

1

PLANNING CHILD HEALTH CARE PROGRAMS:
NEEDS, KNOWLEDGE,
EVALUATION, AND POLICY

by

KATHERINE POWELL MESSENGER
B.A., Smith College
(1970)

SUBMITTED IN
PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF CITY PLANNING

at the

MASSACHUSETTS INSTITUTE OF
TECHNOLOGY

February, 1973

Katherine Powell Messenger
Department of Urban Studies and Planning

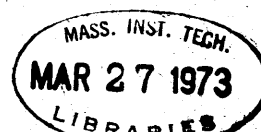
Certified by

Robert M. Hollister
Thesis Advisor

Accepted by

Chairman, Departmental Committee
on Graduate Students

Rotch



To

Sarah

who has done the helping
without needing to ask why

Planning Child Health Care Programs:
Needs, Knowledge, Evaluation and Policy
by
Katherine Powell Messenger

Submitted to the Department of
Urban Studies and Planning
February 7, 1973

in partial fulfillment of the
requirements for the degree of
Master of City Planning

The relationships between evaluation and needs and knowledge, as they are applied to policy issues, are examined in the context of child health care programs. The analysis focuses heavily on a detailed study of social policy evaluation. Definitions of evaluation research, program evaluation, policy analysis and related terms are presented to indicate the lack of conceptual clarity in the field. Goal-attainment and systems models of evaluation theory are defined and the assumptions underlying them analyzed.

The success of evaluation depends in part on facilitating conditions, here termed an "infrastructure." These essential building blocks, which have both cultural and technological dimensions, include: definition of health, health status indices or indicators, determination of health needs, professional standards of care, and the social/political/economic climate.

In the third chapter, after reviewing some models of evaluation, a tentative framework for health care evaluation is proposed which tries to both synthesize the best portions of existing models and to incorporate directly aspects of program operation often neglected. The framework consists of the following components: organization, process, outcome, impact, advocacy, and actionability. In so far as possible, examples have been drawn from child health care issues.

Chapter Four examines major health needs of children and current knowledge about alleviating those needs through social programs. Very little is in fact known about many of the needs. The review focuses on critical needs which affect groups of children, especially those clustered risks which are associated with poverty, high-risk pregnancies, infant mortality, malnutrition, and a continuum of insult and damage to the child. The implications of this continuum for child development models of intervention are discussed. Finally, federal health service programs for children are described; in general we know too little to say if they are really working or not.

The issues raised seem to come together around a concept of boundaries: of health care, of evaluation, and of scientific knowledge. The limitations of all three, which if not respected can convert them from beneficent, useful, and progressive enterprises into meaningless or detrimental ones, are explored in the final chapter. Greater awareness of these boundaries, both analytically and ethically, is advocated, in order that planning and humanitarian goals may be jointly fostered.

Thesis Supervisor: Robert M. Hollister
Title: Assistant Professor of Urban Studies and Planning

TABLE OF CONTENTS

Chapter 1. INTRODUCTION	6
Needs, Knowledge, Evaluation, and Policy	11
Needs	12
Knowledge	14
Evaluation and Policy	15
Plan of the Thesis	19
Chapter 2. EVALUATION: DEFINITIONS AND PRECONDITIONS	24
Part 1. Definitions of Evaluation	24
Evaluation, Value Judgments, and the Scientific Method	29
Evaluation and the Power to Decide	31
Systems vs. Goal-attainment Evaluation	34
Evaluation and Policy Analysis	39
Conclusion	41
Part 2. The Infrastructure of Evaluation	43
Definition of Health	46
Health Status Index or Indicators	47
Determining Health Needs	53
Professional Standards	57
Social/Political/Economic Climate	57
Chapter 3. THE EVALUATION PROCESS: EXISTING FRAMEWORKS AND A TENTATIVE MODEL FOR HEALTH SERVICES EVALUATION	62
Some Frameworks of Evaluation	65
A Tentative Model of Health Services Evaluation	72
Organization	74
Process	77
Outcome	82
Institute of Medicine Methodology of Evaluation	86
Outcome Evaluation and Child Health Care	89
Impact	92
Advocacy	97
Actionability	101

Chapter 4. NEEDS AND KNOWLEDGE: CRITICAL CHILD HEALTH PROBLEMS AND EXISTING APPROACHES TO SOLVE THEM	107
Part 1. Health Needs of Children	107
Child Development and Child Health	108
Normal Patterns of Child Health	116
Critical Factors for Child Health	119
Infant Mortality: Its Causes and Correlates	126
Childhood Mortality	130
Childhood Morbidity	131
Maternal Condition and Risk of Reproductive Complications	135
Nutrition and Child Health	137
Concluding Remarks	149
Part 2. Impact of the Health Care System	150
Access to Care	153
Geographic Access	153
Financial Access	155
Psychological Access	158
Legal Access	161
Special Groups	162
Appropriateness of Care	164
Medical Science	165
Manpower	166
Delivery Modes	168
Quality of Care	169
Part 3. Major Federal Health Service Programs for Children	171
Maternal and Infant Care Program	173
Children and Youth Projects	177
Early Screening and Treatment	179
Family Health Programs	184
Nutritional Programs	187
School Feeding Programs	191
Medicaid and Financing Plans	197
Other Federal Health Programs	199
Chapter 5. CONCLUSION	200
BIBLIOGRAPHY	217

LIST OF TABLES

1. Independent Health Risks 120-121
2. Clustered Health Risks 122-123
3. Estimated Health Care Expenditures by Population and Income
Groups 152
4. Federal Nutrition Programs Affecting Children 188

Chapter One. Introduction

Long ago Machiavelli observed that one cannot be both a good Christian and a good prince. In our secular and democratic society the dilemma persists--can one be both a humanitarian and a good planner? This thesis applies the question to one specific area of social policy: planning health programs for children. I do not know if Machiavelli's cynicism is true today; I do know that the problem is serious, and that it deserves more attention than it now receives.

The heart of the modern conflict between moral action and planning lies in our increasing faith in the power of reason to analyze and solve complex social problems such as poverty, ill-health, and poor education. More precisely, we have faith in technical competence and objectivity to solve these problems. We often forget that these terms are not synonymous. "Reason," which is essentially the core of our humanness and of our commonsensical ways of communicating with each other, has been gradually confined to mean "rationality," a term connoting belief in the superiority of reason over all other means of knowing, and finally to mean "scientific rationality," an even more restricted concept originally appropriate only to matters of scientific inquiry and technology.

The scientific method of research, hypothesis testing, validation, and replication has been gradually but steadily incorporated into first the theories and then the practice of the social sciences. Today, social research methodology and policy science are well-established and respected fields of academic and professional endeavor.

They meet and achieve crucial dimensions in social policy evaluation, a term that, however ambiguous, implies a rational and usually scientifically demonstrable process of analyzing social institutions and programs in order to improve their future performance. Being a good planner implies having a certain amount of faith in that rationality and methodology. Indeed the whole profession of planning is predicated on concepts of the rational determination of problems, the conscious consideration of alternative solutions and the implementation of orderly, planned change.

Yet neither democracy nor social planning is an entirely pragmatic proposition. There are "self-evident truths," "inalienable rights," necessities of life, basic human dignity and worth, and a sincere desire to improve the quality of life--for us and our posterity. Attitudes toward children reflect this humanitarian, moral belief. We value children in a special way as the future of the polity and the society. Not yet capable of full democratic participation or independence, they require special protection and nurturance in order to grow up to be free, healthy, and responsible citizens.

A heterogeneous populace, however, may not have clear consensual notions about the meaning of health, responsibility, or freedom. Functional definitions of these terms must be negotiated. As size makes town meetings and even representative government unwieldy for the bulk of day-to-day policy decisions, the interplay between rationality and ethics becomes more complex. Legislative intent, for example, may indicate that the goal of an act is to better the health and nutrition

of pregnant women, young mothers and infants. In order to judge the effectiveness of the legislation, and to monitor projects arising out of it, bureaucratic guidelines and regulations will "translate" these qualitative goals into less subjective, quantifiable objectives (e.g., to reduce rates of infant and maternal mortality over five years, to train pediatric nurse practitioners or nurse midwives, to provide comprehensive medical and nutritional care to certain groups of women and children, etc.). These objectives are finally combined, constrained by budgetary factors and a desire "to get the most for our money," into cost-effectiveness or cost-benefit measures. And suddenly the self-evident truth of the inalienable right of the infant to life has a price tag on it. We want to guarantee the unalienable right, but we want even more to do it at the lowest cost. We try to show that, while it costs so much to save a child, if healthy that child will eventually generate a certain amount of income, GNP, and purchasing power, to the benefit of us all.

Do we feel concern for children and want to help them because they cannot themselves protect their rights to good health, freedom from hunger, and protection from environmental risks? Or do we act on their behalf because prevention is better (and cheaper) than cure; because healthy children will not require special schools, rehabilitative centers, welfare, hospitalization for long stretches, etc.; and because they will be able to serve in the armed forces, work full time, and pay their own way? Clearly both motivations move us; but the ways in which they are intermingled and the methods by which we reconcile them are not self-apparent. In fact, we need to ask if the values behind them are in basic conflict.

In a book detailing the disappointments and troubles of certain community action poverty programs of the 60's, Marris and Rein capture the essence of this uneasy intermingling by characterizing the "dilemmas of social reform" as the creation of "strategies of action at once politically viable, radically democratic, and scientifically rational." (Marris and Rein, 1969, p. 9).

The political climate has altered dramatically since the heyday of buoyant and reckless optimism of the early War on Poverty. Radical democracy sometimes seems more feared than the old images of China and Russia. But concern with scientific rationality in social policy is not in the least abated. Unlike the New Deal, which was planned by economists (Rossi & Williams, 1972, pp. 12-13), the War on Poverty had battalions of all types of social scientists--psychologists, sociologists, urban planners, policy analysts, operations researchers, etc.--with postwar weapons ready to try social evaluation research, quasi-experimental designs, advanced statistical techniques, survey methodologies, theories of social intervention, etc. Behind all of the jargons and formulations was an underlying confidence that science, social science, planning and political theory could be used in powerful combination to generate policies and programs directly responsive to the most urgent social priorities, to target those policies at the people and places most in need, and to unequivocally demonstrate to disbelievers and to budget-makers the effectiveness and benefits of the policies. Unlike the New Deal, this era of social reform consciously has what Marris and Rein term an "argument":

Reform needs an argument, as well as political and democratic support. In a secular society, where religion is treated as a matter of private conscience, morality is an uneasy ground from which to argue any specific proposal. The sense of shared beliefs begins to dissolve in the light of intellectual analysis. Even though some moral commitment must lie at the foundation of any social policy, it is wiser to pursue, as long as possible, a more dispassionate line of argument. A technologically sophisticated culture promotes, besides, the prestige of scientific problem-solving. Analysis, expert planning, experiment begin to be as much a convention of social policy, as of industrial development or defence strategy. (Marris and Rein, 1969, p. 9)

Thus the enabling legislation of virtually all the health, education, welfare, and community development legislation of the Sixties mandated evaluation and close monitoring both to assure compliance with legislative intent and to measure and document the real-life effectiveness of these previously untried strategies of social reform.*

The history of this new interest in evaluation and effectiveness studies cannot be fully documented here. Among the most significant highpoints were the direct earmarking of program funds for evaluation; the establishment of departmental evaluation, policy analysis and program planning staffs; and especially, the extension of the "Planning-Programming-Budgeting" system (PPBS) from the Department of Defense, where McNamara had used it to completely revamp and trim the defense budget, to the social policy departments (HUD, HEW, and Labor). PPBS, a systems-analysis methodology tying together program planning and budgeting through cost-benefit analysis, is a vastly complex phenomenon

*This is not to imply that specific strategies such as neighborhood health centers, manpower training, or educational reforms had not been tried before; but the social science arguments supporting them were new and more self-conscious. See, for example, Stoekle and Candib's (1969) article on the history of the neighborhood health center as a reform movement.

which cannot be dealt with here. However, as will be evident in some of the discussion which follows, it represents to many people the current best attempt at achieving rational social policy. (c.f. Rivlin, 1970, passim, for probably the best and least excessive formulation of that position; for general history of PPBS and descriptions of its use in the federal government, see Lyden & Miller, 1972; Iglehart, 1972; Rossi, in Rossi & Williams, 1972; and the Proxmire Joint Economic Committee hearings on the subject, U. S. Congress, 1969.)

Concern with "systematic thinking for social action" (Rivlin's title) is increasingly being linked with management science, as the recent appointment of Caspar Weinberger as Secretary of HEW testifies. The question raised by these developments is: to what extent can we transfer to social policy the techniques and assumptions developed to sell products and to manage large-scale enterprises whose purposes are not predicated on evidence of human need but on market demand, profitability, and competition?

I should stress here that I am not trying to make a neo-Luddite, anti-sciences, anti-rational, anti-systems argument. Even the best of ethical and humanitarian programs are better if they are administered fairly, efficiently, and with accountability. But the confidence frequently expressed that scientific methods will eventually remove the messiness and uncertainty from hard choices about policy and human welfare is frightening to me in its hubris.

Needs, Knowledge, Evaluation, and Policy

The complex interplay between desires for social reforms and

scientifically rational ways of achieving those reforms will be examined in the context of child health programs. I have broken down that context into four parts, which provide a structural framework for the analysis to follow. These parts, and the questions which they highlight are presented below.

Needs

Apparent or discovered needs are the basis of social policy. Traditionally needs come to attention through scandal, investigations by individuals (Harrington, Rachel Carson, Upton Sinclair), publicity by political figures (RFK & hunger), public demands (Welfare Rights, civil rights), the publicized success of new ideas (free schools, midwifery, etc.), and other episodic, individualistic, and non-systematic occurrences. They also surface as effective intervention methods are developed to deal with "needs" previously not defined as pressing because there was no solution. Need implies inequality of some sort. When we all suffer without an adequate cure (e.g., from the common cold), or when the risk of susceptibility is apparently random (e.g., many types of cancer), then we are all equal and can be defined in need only in reference to some ideal of health. But, once there were vaccines for polio and rubella, once renal diseases could be controlled via organ transplants and dialysis machines, once amniocentesis could detect inborn metabolic and hereditary errors and made genetic counselling possible, or once mass screening techniques were possible for sickle cell anemia, certain cancers and TB--then we could define those in need as those potential users or beneficiaries of the services who do not receive them or have access to them. Need can also be defined in terms of the

resources necessary on the supply side to meet this demand need (e.g., organs for transplantation, health personnel to staff screening centers, legal clarifications on rights to abortion, euthanasia).

Finally, need may be determined deliberately and synthetically as existing social institutions are examined (evaluated) in order to determine whether they are performing as common knowledge says they are and whether there are unnoticed gaps or inequities being perpetrated. Thus the Coleman report and its apparently endless offspring have created a self-conscious examination of the assumptions behind universal public education and compensatory education. Similarly, the current interest in HMO's and other models of organized, prepaid group health care arises from longstanding, but only recently widely acknowledged, challenges to the tenets of private fee-for-service medicine and from experiments with alternate schemes. In this sense, the "crisis in health care" is indeed a manufactured one. The same analytic studies and conclusions were reached 40 years ago by the Committee on the Costs of Medical Care for the American People (1927-32).

Clearly the definitions outlined above are overlapping. But they do suggest that the proportions of passion and analysis in the determination of need are quite variable. In the end, as I shall discuss in the next chapter, scientific rationality (via tools like epidemiology, statistical procedures, and surveys) must in some way quantify or at least objectify even the most impassioned cries of need. How many kidney machines are enough? How many migrant children never get adequate medical care? What sorts of services do a handicapped child and his family need?

Question: What do we know, from all of these sources, about the health needs of American children?

Question: Are there ways in which social needs can be more systematically reviewed and assessed, without concentrating on those which are merely more obvious or more easily quantified?

Knowledge

Knowledge here refers to:

1) Specific responses effective to meet certain needs. These generally arise from small-scale, demonstration, or experimental projects or research. In the health field, such responses range from medical techniques and drugs to strategies of service delivery and intervention. They range also from well-tested and verified knowledge (e.g., drug efficacy after human trials) to questionable extensions of theory into practice and untested transfers from allegedly comparable situations, and to traditional "time-tested" responses (e.g., private physicians).

2) Theories of social behavior or medical etiology. These usually form the basis of proposals to meet wider sets of needs. For example, theories of familial and environmental stress as causal factors in mental illness influenced the "argument" for community mental health and preventive psychiatry; similar theories now influence new broadly-based approaches to child abuse and neglect. Here the important role of knowledge is to resynthesize or reorganize isolated pieces of intervention strategies into new matrices. While the parts are important, the whole is the heart of the argument. These matrices can range from new approaches to the organization of hospitals to such massive programs as the community action ones of the War on Poverty.

(Model Cities, for example, was supposed to reverse urban decay, social disintegration, and racial tensions.) In health, important examples are the arguments linking good prenatal and infant care with cognitive and physical development and school performance; arguments for family-oriented and organized health care; and ecological models of the relationship between the child's health and his environment (accidents, lead paint poisoning, child abuse, malnutrition, rat bites, family stress, etc.).

3) Finally, there are the overarching theories, the dominant modes of analysis in the society, those key concepts and beliefs which almost unconsciously influence behavior and policy. "Child development" is such a model of particular interest here. The old Children's Bureau is dead, the Office of Child Development its reincarnation. This is more than renaming; the focus is different, the stress on cognitive development, stages (à la Piaget and Erikson), and calculated intervention. Much of HEW and OEO's research money is going into child development-oriented inquiries.

Question: What do we know about programs for child health? What kinds of things are possible? What models of child health are useful?

Question: How do health and health care relate to child development? Is this latter concept useful or limiting in structuring strategies to meet the health needs of children?

Evaluation and Policy

Public policy ideally should be the application of knowledge of many kinds to the most pressing needs of the society. The activities

associated with it are synthesizing and reviewing data on needs and knowledge, making priority rankings, creating programs or strategies to bring about desired changes, and implementing these strategies. Fundamentally it is a process of resource allocation.

Evaluation means many things to many people. The International Encyclopedia of Social Science defines the purpose of evaluation research:

to provide objective, systematic and comprehensive evidence on the degree to which the program achieves its intended objectives plus the degree to which it produces other unanticipated consequences, which when recognized would also be regarded as relevant. (IESS, 1968, vol. 5, p. 198)

Evaluation here means the planned and organized gathering of knowledge about social action. That is, evaluation adds to our store of knowledge about what interventions work and to some extent, about how they work.

Other definitions of evaluation stress its more direct feedback potential into the policy-making apparatus which created the program under study.

In order to influence social policy, findings from social-action experiments must provide a basis for the efficient allocation of financial and human resources in the solution of social problems. It is this notion of the efficient allocation of resources that is the key to the whole problem of planning and choosing among social-action programs. (Freeman & Sherwood, 1965; reprinted in Schulberg et al., 1969, p. 80)

Evaluation here is seen as an integral component of the social policy process, a critical means of assessing priorities in the allocation of resources. The definition also introduces the criterion of efficiency as the key to planning, an assessment not universally shared. Both this definition and the previous one strongly suggest the rationality and scientific methodology of evaluation ("objective,"

"systematic," "evidence," "degree," "experiments," "solution," "efficient allocation").

A third definition, from a study committee on educational evaluation, offers a significantly different perspective:

Evaluation is the process of delineating, obtaining, and providing useful information for judging decision alternatives.
(Stufflebeam et al., 1971, p. 40)

Here evaluation is defined as a broad process providing the basis for making social policy decisions. It does not single out the usual "compare performance with intended objectives" as its main function. As I will discuss in Chapter 3, that function is difficult to perform and of ambiguous meaning, although such evaluation (the IESS definition) would certainly be one kind of "useful information." But "useful" is a non-scientific word which can also encompass many other types of "data." Most importantly, this last definition explicitly acknowledges the subjective nature of evaluation policy-making, that it means "judging," not "planning and choosing." As Weckworth, a prominent evaluator of health programs has said, evaluation "ultimately becomes judgment, i.e. the opinion of the person with the right to decide." (Weckworth, 1969, p. 12). If this is true, then we need to examine carefully how evaluation is done, what types of information are gathered, to whom it is given, and in what types of decisions it is instrumental.

This judgment returns us to Machiavelli. Can the argument of social reform encompass both the scientifically rational, the known, the demonstrated, the measured--and the ethically sound, the toughest of issues, the commonsensically important, the unquantifiable?

According to what criteria is it decided which of the multitude of issues in medical care, biomedical research, psychosocial aspects of health and health care delivery, special needs of special groups, organizational problems, etc. shall be examined? If evaluation, and ultimately policy action, depends on knowledge, certain reforms may be foreclosed because problems are ignored and basic questions never asked.* In a society that is affluent, relatively healthy, and well supplied with medical resources, very few new programs or treatments will result in rapid, startling, massive changes in health status or levels of access.

Social science methodologies of evaluative research are the major evaluation tool; yet they have many serious shortcomings. They are not well developed either to measure and verify rather small changes in complex milieus or to apportion the effect of each of many causal factors on such measures of health outcome as infant mortality, reduced community illness, or family intactness. There are a multitude of statistical techniques available, and reality rarely fully matches the needs of any of them; methodology can always be challenged by critics of the conclusions.

In view of this, how do we decide whether programs are in fact effective or not? Having once entered the arena of evaluative research, experimental designs and scientifically rational arguments, where do we draw the line? If everyone serving and being served feels a program

*For example, sickle cell anemia is not a new disease, nor only recently widespread; but because it afflicts mainly blacks, it has never been much studied by white-dominated NIH or other researchers; hence we did not know much about its extent, treatment or detection.

is useful and of high quality, do we continue to question it if we can't document impressive changes? Why were the revelations of hunger by the Citizens Board of Inquiry discounted because they were only testimony and case studies? How often do we challenge the social scientist's findings of "no significant difference" by saying, "When a technique continually produces findings that are at variance with experience and common observation, it is time to call that technique into question." (Stufflebeam et al., 1971, p. 8) And, if we must have some policy, for health or children, how do we act in the face of incomplete and confusing knowledge while maintaining belief in our ability, and indeed our responsibility, to act reasonably, on the basis of explicable argument?

Plan of the Thesis

In closing this introduction, I should briefly lay out the circumstances under which this thesis came to be in its present shape. Most of the research was done as part of my work for the Huron Institute of Cambridge, Massachusetts on a major project, under contract to the Office of the Assistant Secretary for Planning and Evaluation, HEW, originally entitled "Federal Child Development Programs for Disadvantaged Children." The program analysis, which lasted over a year (Summer/1971 to September/1972) under the direction of Sheldon H. White of Harvard, was originally designed to systematically review the literature in five areas believed critical to child development: elementary education, preschool education, day care, family intervention, and health. I was responsible for developing the health section of the report. The report was intended to synthesize the literature into critical periods

for intervention and sets of critical needs of disadvantaged children (having first developed a comprehensive definition of disadvantage); to review existing evaluative tools and standards of measurement; to survey findings from existing programs (federal and non-federal); to then combine these sets of data so that we could derive conclusive recommendations for future policy (short-, intermediate-, and long-term) based on critical needs, periods, and performance evidence; and finally, to cost out a range of these recommendations.

The project was a retrenchment from the overblown, untried theories of the Sixties, a rigorously scientific and methodologically self-conscious enterprise, infused with the current interest in child development and hoping to establish the latter's validity as an operational policy construct, one with much greater power and truth than the more political or vaguely humanitarian arguments of the Great Society. There was a real belief at the beginning that if only we studied it carefully and comprehensively enough, the literature would reveal solid knowledge and unarguable bases for action. And yet, the overwhelming--and to many, deeply disillusioning--finding was that we could recommend virtually nothing (and certainly nothing major or comprehensive) on the basis of conclusive evidence. And we could certainly not cost out ranges of possible programs. But the project (retitled "Federal Programs for Young Children") did come up with recommendations--150 pages of them. There was little evidence, we decided, that we should not attempt intervention in processes of child development, health, or education, and therefore recommendations were developed inductively from experience, testimony, trends in data, pieces of evidence, and values--and then

buttressed by organized presentations of the data previously amassed.

I found myself torn between wanting to argue the cases on unassailable grounds of "hard data" (particularly since I had a sense of writing for a basically hostile or at least budget-cutting administration), but also wanting not to let go of deeply held beliefs (such as the importance of consumer control or of services as rights rather than as welfare) simply because they did not fit the organizing framework of child development theories or lacked evidence of previous success. The tensions created by the situation were strong enough to make me reject the validity of any of the work I had done for several months afterwards. Finally I realized that the complexities of the issues raised, my own dissatisfaction with either the conclusions or the whole process of the work went beyond the immediate situation or my own despair of doing better. This thesis, then, represents a first attempt on my part to apply my own analysis and value judgments to the questions not addressed in that report: how do we go about deciding what is knowledge and how do we use it or act on the basis of it.

As I have presented them, needs, knowledge, evaluation and policy are a looped system which can be viewed as beginning with any one of the components. For my purposes, evaluation is central both as an organizing framework and as a constraint upon the dimensions of the other components. For these reasons, I am devoting the bulk of this thesis to an examination of evaluation in health care programs.

Chapter Two will present a range of definitions of evaluation, discuss the most common assumptions underlying these definitions and

theories, and explore a set of prerequisite conditions (termed an "infrastructure") necessary for meaningful evaluations to be carried out. Reference will be made specifically to definitions and conditions for health program evaluation and to issues of direct impact on children whenever possible, although more general observations will often be necessary.

Chapter Three deals with the content and process of evaluation itself. The first part of the chapter presents current models, theories, and frameworks of evaluation (primarily designed specifically for health evaluation) to suggest the range of variables considered relevant for evaluation methodologies. The second part of the chapter is a tentative framework for evaluating health programs which I have developed by synthesizing existing models and incorporating some aspects of concern which they frequently ignore. The purpose of these two chapters is both to reveal the confused state of theories of evaluations and the even more dismal general record of practice, and to provide evidence for my conviction that better practice and theory are possible given the constraints of current tools and beliefs.

In Chapter Four, I return to the specific issue of child health, to explore the outlines of our knowledge about the needs of children and about the effects of existing federal programs on child health. The origins of concern in preventive medicine, child development theories, and in the unfinished business of the Sixties will be briefly presented. In looking at the critical health needs of children--both individually and as users (or nonusers) of health care services--I stress epidemiological, sociological, and politico-economic issues rather than those of basic

medical research or expertise. Hence the discussion slights specific treatment needs but goes into depth on the effects of the health care system itself on the health chances of children. The final part of the chapter summarizes findings from some major federal health service programs for children, examining both the methods by which they have been evaluated and the evidence they have generated on useful and proper intervention techniques. Because the number of such evaluations is small and the number of well-designed or broadly significant ones smaller still, the discussion will be supplemented with evidence (both clinical and programmatic) and testimony from some non-federal efforts to improve child health. This is not a comprehensive review of efforts in child health and related fields but rather an example of the types of information available for policy decisions and of potentially useful ways of organizing it to feed back into policy and planning processes.

Policy and planning methods are not dealt with directly in the thesis. Originally, a fifth chapter was to present several examples of policy recommendations based on sets of evaluation findings, to study the assumptions under which these recommendations were made, and to evaluate the validity or the scientific rationality of those decisions. This has proven to be a topic worthy of a lengthy paper in its own right. Throughout this paper, however, issues of policy are constantly raised and related to the issues of evaluation methodology, highlighting the complexities of the issues and the dangers of seeking quick or simple answers to them.

A final chapter draws some tentative conclusions about the relationship of values to knowledge in evaluation and planning, assesses the extent to which the questions raised in this Introduction have been answered, and raises some further questions still unanswered.

Chapter 2.

Evaluation: Definitions and Preconditions

This chapter has a dual purpose. First, it will present current theoretical and programmatic definitions of evaluation. Second, it will examine the contextual assumptions behind these definitions and will discuss in detail the preconditions required for those assumptions to be met.

Part 1. Definitions of Evaluation

In reviewing the literature on social policy evaluation, one is immediately struck by the fact that there is virtually no agreement on even basic definitions or terminology. A certain amount of "saying things your own way" and creating new categories is to be expected just to add the appearance, or reality, of originality to one's work, to correct the less developed theories of the past. But many writers on the subject of evaluation research or evaluation of social programs do not seem aware of each other's writings. Because they have not been brought together before, the collection of definitions below, which might otherwise be a rather pedestrian, routine task, has proven to be both time-consuming and revealing.

To begin, there is one clearly defined concept of evaluation research, the only entry related to evaluation in the International Encyclopedia of the Social Sciences. (1968) That definition reads: "A scientific approach to the assessment of a program's achievements is the hallmark of modern research evaluation...The primary purpose of evaluation

is 'to provide objective, systematic, and comprehensive evidence on the degree to which it produces other unanticipated consequences, which when recognized would also be regarded as relevant to the agency'". (IESS, 1968, Vol. 5, p. 198) Evaluation research is then contrasted with other types of social research, such as exploratory studies to formulate new problems and hypotheses (corresponding to my use of the word "knowledge" as learning about needs), explanatory research, and descriptive social research. Inherent in this definition are the following five methodological steps:

- 1) Conceptualization and measurement of objectives of the program and other unanticipated relevant outcomes;
- 2) Formulation of a research design and the criteria for proof of effectiveness of the program, including consideration of control groups or alternatives to them;
- 3) The development and application of research procedures, including provisions for the estimation or reduction of errors in measurement;
- 4) Problems of index construction and the proper evaluation of effectiveness; and
- 5) Procedures for understanding and explaining the findings on effectiveness or ineffectiveness. (IESS, Vol. 5, p. 199)

Evaluation research is similarly defined by Suchman, probably the single most influential and important writer in the field, as "the use of the scientific method for collecting data concerning the degree to which some specified activity achieves some desired effect." (Suchman, 1969, p. 15) Suchman differentiates it from evaluation which he defines as a "general social process of making judgments of worth regardless of the basis for such judgments," and from basic research because it is applied to problems with administrative consequences rather than theoretical significance. (c.f. Caro, 1971, p. 8; James, 1962; similar definitions and distinctions are common in the mental health field, c.f.

NIMH; 1955, MacMahon, Pugh & Hutchison, 1961; Edgerton, 1971)

In a discussion of evaluation at the federal level, Wholey et al. (1970) define evaluation operationally as that process which

- 1) assesses the effectiveness of an on-going program in achieving its objectives,
- 2) relies on the principles of research design to distinguish a program's effects from those of other forces working in a situation, and
- 3) aims at program improvement through a modification of current operations. (Wholey, et al., 1970, p. 23; emphasis in original)

The authors contrast evaluation with program efficiency studies, which focuses on input rather than output or goals as evaluation does, and with program analysis and policy analysis which usually compare both existing and hypothetical programs designed to solve the same problems. They further note two alternatives to evaluation for gaining information about on-going programs: field experiments (designed with control groups) and experimental demonstrations (designed without control groups).

Wholey himself, in a later piece, uses the term program evaluation in an apparently analogous manner, defining it as "assessment of the impact of past and present programs, projects, and project components ('treatments')." (Wholey, undated, p. I.1)

For a definition directly related to health programs, one can use the American Public Health Association's definition of evaluation as the "process of determining the value or amount of success in achieving a predetermined objective." (APHA, 1960) Schulberg, Sheldon, and Baker (1969), in the introduction to their reader on evaluation in the health fields, state that this is the sense in which "program evaluation" is usually used and that it has the following four steps inherent in it (c.f. steps noted above in IESS definition):

- 1) formulating the objective;
- 2) identifying proper criteria to be used in measuring success;
- 3) determining and explaining the program's degree of success;
- 4) and recommending further program activity.

The authors at this point also attempt a definition of program, a central term rarely discussed. They say:

In the field of organizational study, programs generally are defined as a set of activities occurring within a social enterprise which have specific inputs of resources and conditions, techniques for establishing relations among them, and certain outputs which can be evaluated against given standards. Additionally, aspects of the organization's patterned activities occur not only within its own structure but also in relation to other organizations as well. (p. 5)

We will return later to the immense difficulty, which Schulberg et al. go on to discuss, of actually finding health programs with such clearcut boundaries, goals, and objectives.

Schulberg et al. show that some authors distinguish program evaluation from other research by its attempts to mark progress toward prestated goals instead of seeking new knowledge (c.f. IESS's "exploratory studies" and Suchman's "basic research"; also James, 1962; Edgerton, 1971), and from program reviews or demonstrations, by its more rigorous methodological requirements. They make an important comment on the arbitrary nature of some of these distinctions.

The information obtained by assessing a program's ability, or inability, to meet prestated objectives can be useful not only for pragmatic reasons but also for generating new knowledge about the disease entity being treated. (Schulberg, et al., 1969, p.6)

To summarize to this point: there seems to be some basic distinction being made among three levels of endeavor. First, there is something called basic research, be it medical, sociological, or psychological, which examines questions of knowledge which are not tied directly to the operations of any human services or social action program according to scientific, experimental

methodologies. Second, there is something which we will call evaluation research which meets the IESS, Suchman, and APHA definitions of using experimental research methodology to assess the effectiveness of a program in meeting stated goals or objectives. The IESS definition is the only one of these which also requires evaluation research to consciously look for unanticipated effects.* Third, there is what we will call program evaluation, which focuses on program change as the basis of evaluation research. This administrative feedback mechanism is explicit in the definition of Wholey et al., in the steps which Schulberg et al. tie to the APHA definition, and in the more narrow distinction drawn by two evaluators of the OEO Neighborhood Health Centers:

While program evaluation draws on the technology of research, it is clearly distinguished from research by its relevance to annual policy, legislative and budgetary concerns, and to program planning. (Sparer and Johnson, 1970, p. 4)

One can push the distinctions between these categories only so far. Qualitatively there is an obvious difference in level of analysis and purpose between an academic research study on the effects of a new drug for renal disease involving 15 elderly patients, and an HEW study comparing rates of infant mortality and levels of infant health among all the projects of a federal Maternal and Infant Care program, due for completion in time for annual budget review procedures. But both studies influence both our store of knowledge and the eventual path of social action programs. Succinctly put, the critical characteristic of both evaluation research and program evaluations "is that decisions as to resource allocation or program change are likely to hinge on [their] outcomes." (Edgerton, 1971, p. 94)

* Although, since the "unanticipated" results must be anticipated in designing the methodology, this distinction may be rather arbitrary and reflect only a priority ranking of the importance of expected effects or of those of greatest policy interest.

Distinctions between evaluation research and program evaluation stem mainly from the background and professional positions of the evaluators, not from the nature of the evaluation itself. The term "program evaluation" seems to be used more frequently by those in positions of administering or studying a variety of programs aimed at the same problems (e.g., federal bureaucrats, mental health or public health department administrators, or academic researchers interested in a field of social intervention, such as maternal and infant care, community mental health or health education). In the remainder of the paper, we shall use evaluation research and program evaluation interchangeably, unless otherwise noted in reference to the terminology of a specific author.

Evaluation, Value Judgments and the Scientific Method

All of the definitions and methodologies discussed above refer to or implicitly suggest the importance of the scientific method (i.e. experimental design) in the conduct of evaluation. They reflect one of the origins of evaluation, in social science research methodology. Viewed in this way, evaluation of human service programs is the culmination of a slow extension of experimental methods from physical science to biological science to economics, psychology, and to the other social sciences (sociology, political science, and planning). It reflects a growing competence and confidence in understanding reality by explicitly testing hypotheses and observing outcomes in complex social situations as well as in refined, highly controlled laboratory settings.

Evaluation, however, can also be viewed as part of the most recent paradigm of organizational and decision-making theory. This history also involves many of the disciplines mentioned above, but did not originate

entirely in the "hard" sciences. Thus it has never placed primary emphasis on scientific or experimental methods, but rather has sought a wide variety of explanatory modes to analyze how social organizations begin, define goals, operate, change, and make decisions. It emphasizes interpersonal relations as much as it does game theory and computerized decision-making trees. From this perspective, there are a conceptually completely different set of definitions of evaluation. It is the set which Suchman terms evaluation, "a general social process of making judgments of worth regardless of the basis for such judgments." (Suchman, 1969) Similarly, a Study Committee on Educational Evaluation defined evaluation as "the process of delineating, obtaining, and providing useful information for judging decision alternatives." (Stufflebeam et al., 1971, p. 40) In their discussion, which we shall examine in greater detail below, they make clear that this definition deliberately extends the parameters of "useful information" beyond the usual research methodologies, measurement criteria, and comparisons of performance with stated objectives.

The two sets of definitions suggest a crucial qualitative distinction directly related to a conception of the dichotomy between social value or humanitarian grounds of argument and scientifically rational ones. It is not trivial that some evaluators define evaluation research as requiring a rigorous methodology, and then note a residual category, suggestively untrustworthy ("regardless of the basis"), while others define evaluation in a broad, qualitative way with scientific methodology evaluation being a sub-category. These latter definitions recognize explicitly that evaluation (in any formulation) "ultimately becomes judgment, i.e. the opinion of the person with the right to decide." (Weckworth, 1969, p. 12) The dictionary definition of evaluation is "to determine or fix a value of"

or "to examine and judge". There are many ways in which this power can be exercised: by deciding which social activities will be evaluated, by setting the ground rules for the evaluation (who will be interviewed or observed, whose records studied), by determining the methods of evaluation and the measures of objectives to be used, and by having power over dissemination and utilization of the findings.

In attempts to reconcile the methodologic demands of scientific evaluation research with the realities of political decision-making, at least two approaches seem common. One adds a concern with administrative constraints onto the evaluator's work plan; the other shifts the basic parameters of evaluation away from its traditional goal-oriented thinking. The first augments the previous definitions of good evaluation, the second changes them significantly. A discussion of each follows.

Evaluation and the Power to Decide

Virtually all social policy evaluators who have done any real-life evaluations discuss the problem of implementing the policy recommendations of completed evaluations or indeed of achieving enough administrative trust to do well-designed and thorough evaluations. (c.f. the introductions in Caro, 1971, and Schulberg, Sheldon & Baker, 1969; and Rossi & Williams, 1972 (Chap. I) for excellent summaries of the usual problems and references to appropriate readings). From the viewpoint of both program administrators and proponents of specific social reforms, the problem is that there is "political vulnerability from knowing outcomes." (Campbell, 1969, in Caro, 1971, p. 234) In a direct and personal sense, administrators feel threatened by the possibility of negative or even no-effect findings. This fear is realistic given that most research designs can demonstrate lack of

correlations but cannot prove causal relationships. And methodologies, whether they reveal positive or negative results, are always open to attack by a program's critics. (E.g., the furor raised in OEO and elsewhere by the Westinghouse-Ohio State study of Head Start; c.f. Williams & Evans, 1969; Campbell & Erlebacher, 1970) Evaluation techniques often succeed only in stripping the scientific facade off programs to reveal the value conflicts beneath, value conflicts which evaluation itself cannot resolve. In a society which likes to pretend that it is modern, scientific, and one big happy consensual family, these facades of goals and explanations of intended effects are often vital in permitting progress past (or at least around) value log jams. While we expound scientific rationalism and rules of evidence and proof, we know we can't really wait to act on such bases. When confronted with the possibility of evaluation studies which take scientific rigor seriously and are quite content to measure only those variables capable of quantification, almost everyone but the Office of Management and Budget retreats a little.

The methodological implications of this over-rationality will be discussed later. Here we want to illustrate some of the ways proposed to systematically consider "the consequences of the evaluation and how these will or should affect the evaluation." (Harper & Babigan, 1971, p. 151)

Harper and Babigan (1971) propose that "the range of possible recommendations, that is, those which have some chance of being implemented, should be arrayed before the evaluation and should, in fact, influence the evaluation." (p. 152) While this reeks of "fudging the results" to a strict researcher, it seems both plausible and necessary in policy-oriented work. As the authors illustrate, finding out beforehand that

there are no funds to substantially increase a program even if it is working or that the program will not be defunded no matter what because the governor's sister suffers from the disease it treats means that the evaluator can focus his design on gathering information on the most successful subparts of the program in order to accomplish more with the same money or on how to restructure the program to make it more responsive without destroying it. This certainly doesn't solve any dilemmas for the evaluator if he in fact finds no evidence of impact, but at least he won't waste all his energy beating dead issues.

This approach may sound unnecessarily defeatist until one reads enough sad tales of evaluators who never gained access to key people or records in a program because the administrator wouldn't let them, of studies which never got finished, and even worse of the final studies which were never read by anyone except the next generation of planners or evaluators. Marris and Rein's Dilemmas of Social Reform (1969) provides an excellent insightful discussion of some of these problems in the community action programs of the 60's. (See also Weiss & Rein, 1969; and Section V., "Implementing Research Findings", passim, in Schulberg, Sheldon & Baker, 1969)

Putting the same issue in a more positive light, Rossi (Chap. 2 in Rossi & Williams, 1972) argues strongly that evaluators have an obligation to "pay some attention" to the outcomes of their activities. He calls this duty a "commitment to evaluation".

By commitment to evaluation is meant that it should be worked out in advance what are the policy changes that will flow from each of the set of possible findings...Commitment does not necessarily mean that evaluation research should take the place of the traditional decision-making process, but it does mean that at least the parties involved will have thought through in advance how they might respond under various likely contingencies. (Rossi, in Rossi & Williams, 1972, p. 45)

Freeman and Sherwood (1965) note that demands for evaluation of major social programs are frequently only a formality and studies initiated are often left undone.

As a consequence, adequately conceived efforts have in fact been undertaken rarely and the sheer infrequency of completed investigation is a major reason for the minimal impact of evaluation research on social policy. Certainly it is difficult to point to many instances in which programs actually have been modified, expanded or terminated because of evaluation findings. (In Schulberg, et al., p.74)

Their definition of evaluation cited in the Introduction stresses that its link to social policy is in providing a basis for "the efficient allocation of financial and human resources in the solution of social problems." (Ibid, p.80) One important parameter of that allocation process is knowledge of acceptable outer bounds for the choices.

Finally, a study committee on evaluation has noted that one consequence of the lack of an adequate theory of evaluation is a concomitant lack of:

any specification of the kind of evaluative information which could be most useful. Since the amount and kind of information which could be collected is infinitely large, this lack can be crucial. (Stufflebeam, et al., 1971, p.8)

A closer, more discriminating consideration of possible evaluation findings and implementations, and the information most useful for each is perhaps the most reasonable response to the impossibility of achieving complete knowledge or apolitical policy-making.

Systems vs. Goal-Attainment Evaluation

Some critics of typical evaluation research feel that the addition of techniques for dealing with feedback into policy and with administrative hostility is not the best answer to the evident limitations of evaluation. They believe that there is a more fundamental reason why

evaluation does not "work" very often. Evaluation designs which direct their attention solely to the comparison of performance with stated objectives have, no matter how complex and sensitive the process, a false conception of how programs work.

The alternative usually proposed is a systems orientation to replace the typical goal-attainment one. Some generally excellent analysts of evaluation, such as Suchman (1969), never even mention systems evaluation. The fullest comparison of the two approaches I found was Schulberg, Sheldon and Baker (1969). All of the definitions presented so far are variations on the goal-attainment model. In contrast, proponents of a systems model, of which Amitai Etzioni is the foremost, feel that the starting point for evaluation should not be a specific goal or objective, "but rather a working model of the social unit which is capable of achieving a goal". (Etzioni, 1960, in Schulberg, et al., 1969, p. 10) Since real life systems must devote resources to functions other than goal-attainment, this framework enables one to ask the basic question:

"How close does the organization's allocation of resources approach an optimum distribution?" (Ibid.)

In other words, needs and functions such as service and custodial activities need to be balanced against goal-achievement activities. In a manner analogous to economic resource allocation models, there may be a "diminishing return on investment" of focusing too much effort on a single organizational goal. Intuitively this reasoning is obvious; but it is not always recognized in practice that evaluating only one aspect of performance will not provide answers as to whether that performance is optimal for that organization, or for the country.

Systems models of evaluation emphasize the multipurpose nature of health care institutions or programs and the organizational context both of service delivery and of evaluation.* This later consideration is basically the same as the issues of implementation and feedback raised in the last section.

Among other reasons, Etzioni (1960) gives the two following explanations of why the goal model may not be the best frame for judging effectiveness of social programs. First:

It compares the ideal with the real, as a result of which most levels of performance look alike - quite low. (In Schulberg, et al, 1969, p. 103)

This consideration is important both because research methodology is not very good at measuring and verifying small changes in outcomes and because, as people like Dubos (1959) have shown, the days of major dramatic breakthroughs in health due to single program efforts may be gone. The predominance of chronic diseases and mental illness over infectious, acute diseases, the universal leveling off of the infant mortality curves at a residual rate in all advanced countries, and the increasing problems with iatrogenic illnesses all lend support to the notion that only trivial objectives may show great performance outcomes. In discussing the obligation of evaluators to be aware of the impact of their studies, Rossi has noted:

It is part of our responsibility as social scientists and as researchers to make everyone aware that in this period even the best of social action programs are not likely to produce spectacular results. The age of miracles is long over. New programs can be expected when they are success-

* Program impact evaluations which compare the success of several projects in the same program may be able to do this if the projects are administratively and contextually similar.

ful to be only somewhat better than existing programs. Any new program is not likely to produce spectacular changes, and a good proportion are likely to fail to produce any detectable changes. (Rossi, in Rossi & Williams, 1972, p. 45)

Second, Etzioni warns against too great an emphasis on pre-stated goals, especially public ones:

Public goals fail to be realized not because of poor planning, unanticipated consequences, or hostile environment. They are not meant to be realized. (Ibid, emphasis in original)

Being used less frequently in social program evaluations, methodologies applicable to the systems model are less agreed upon. Basically they are those which are able to handle sets ^{of} variables, all of which are dependent and reciprocal, in true feedback systems constructs. Techniques such as multivariate analysis and path analysis are appropriate. There are problems, however, both with the complex, expensive and time-consuming nature of these techniques and with the limits of such an ambitious approach to knowledge.

In contrast to the goal-attainment model which simply requires the researcher specify the particular organizational goal which he wishes to study, the systems model requires the evaluator to determine what he considers a highly effective allocation of means and then to study the organization's degree of success in achieving this optimal distribution. (Schulberg, Sheldon & Baker, 1969, p. 11)

I am of divided mind about the importance of systems evaluation. On the one hand, from a limited knowledge of organizational theory and its applications, I do not have the impression that it is really more useful for making policy decisions (as opposed to providing an awareness of the complex context of those decisions) than more limited evaluations which test specific service/outcome linkages. Too much attention to the many other functions of organizations and to the enormous homeostatic and inertial tendencies they possess may bias

the evaluator too early as to the impossibility of change.* Furthermore, without clear prior knowledge about those organizational patterns accessible to change, the evaluator may still be faced with interesting and well-documented recommendations about non-policy variables.

On the other hand, systems evaluation designs which can accommodate multi-objective programs and their contextual relationships may be particularly appropriate for free-standing, innovative programs where the whole process proposed is part of the rationale of the venture. For example, neighborhood health centers, maternal and infant care projects, freely accessible abortion services, or even prepaid group practice are qualitatively different from redesigned emergency rooms, introduction of a new screening procedure for diabetes, or expanded availability of food stamps as health policy decisions. The latter are additions to or incremental changes in existing programs and systems of health and nutrition services while the former are in effect new systems and approaches to health care. Because of the overlap between these categories neither model of evaluation should be used exclusively, but the pay-off from the additional effort required for systems evaluation should be far greater for the former programs. The only systems-model study of a current major federal health program

* It would seem axiomatic that those values underlying any so-called "value-free" theory of social functioning reveal the unexamined assumptions of that theory. It is no accident that Parsonian structural-functional analysis is not a revolutionary mode of analysis or that traditional micro-economic theory views conglomerates and the military-industrial complex as aberrations of the capitalist system, while Marxist or political economic analysis begin with their inevitability. Thus it is worthy of comment that Etzioni feels obliged to argue rather defensively why his model is not inherently conservative. (In Schulberg, 1969)

I am aware of is the Geomet evaluation of Neighborhood Health Centers (Geomet, 1971) which not only examines the interactions between project structure and process and outcome measures but also makes a serious attempt to develop a conceptual framework for evaluating the effects of NHC's on poverty, one of the "public goals" of the programs.

Evaluation and Policy Analysis

To complete this anthology of definitions, policy analysis should be differentiated from evaluation research. As noted in the Introduction, this distinction is clearest at the federal level. Walter Williams, one of the best known writers on the application of social policy evaluation to government programs, has defined policy analysis as

a policy-oriented approach, method, and collection of techniques of synthesizing information including the results of research: (a) to specify alternative policy and program choices and preferred alternatives in comparable, predicted qualitative and quantitative cost/benefit type terms as a format for decision making; (b) to assess organizational goals in terms of value inputs and to specify the requisite output criteria for organizational goals as a basis of goal determination and measurement of outcome performance; and (c) to determine needed additional information in support of policy analysis as a guide for future decisions concerning analytical and research activities. (Williams, 1971, pp. 12-13)

Williams distinguishes evaluation research (meaning outcome evaluation and field experiments) by its direct information-gathering, rather than -utilizing, function. The reason for the careful distinction is to highlight the fact that "the results of research are an input to analysis that may limit severely its successful application." (Williams, in Rossi & Williams, 1972, p. 4) Analysts can only stretch incomplete socio-economic data, cost estimates, and program alternatives so far before policy analysis becomes only a pretense. In the years since the initial interest in PPBS and other scientifically rational policy

analysis tools in social policy, optimism in their utility and power has not proved realistic. Williams offers the following explanation:

As time passed it became clear that the critical missing element was research treating specifically issues of program conceptualization, design, operation and measurement. In short, the central analyst in social agencies could not show to policy makers evaluative research results indicating that a current program worked or that an alternative to that program was likely to work better. (Williams, in Rossi & Williams, 1972, p. 5)

The remainder of this chapter and all of the next will be devoted to an analysis of how far evaluation can and cannot be expected to go in providing such results. I will not attempt to analyze policy analysis or cost-benefit analysis in similar detail, but there are other possible explanations for the indifferent success of policy analysis than the inadequacies of evaluation research. Williams' definition (a typical one, by the way) is almost a classical example of absolute faith in rationality and plannerly competence to control environments. Not even by implication are the immense moral and value judgments inherent in the definition acknowledged. Whose "preferred" alternatives? Whose "value inputs"? As if values were something one injected at precise moments and in exact amounts in the production process of policy. How do qualitative matters get translated into "cost benefit type terms"? Perhaps it is unfair to fault an argument on the basis of the quality of the language in which it is couched, but Williams' definition is both inelegant and lacking in substance under close examination. The simplistic definition "compare performance with stated objectives" suggests more of what both the process and product of evaluation might be than this elaborate, but contentless formulation.

Conclusion

I have presented what seem to me the major definitions of evaluation research and evaluation along with some responses to the limitations of those definitions. Some major issues remain unresolved:

- 1) The extent to which the validity of "evaluation" is dependent on or shaped by social science research methodology;
- 2) The emphasis which evaluation places on goal-achievement as a criterion for program effectiveness, the core concern of all the definitions;
- 3) The possibility of making the linkages between evaluation and policy part of evaluation design; and
- 4) The degree to which evaluation is both acknowledged to be and designed as an exercise of human judgment (i.e. value-laden, ethical, or indeed amoral) on the course of social policy.

I will attempt to pay particular attention in the pages to follow to the difference between goal-oriented and systems models of evaluation, since this seems the most concrete and programmatically significant distinction drawn in the definitions. However, because of the preponderance of goal-attainment models in both theory and practice, most of my examples have had to be drawn from that school. Hence, a synopsis of the assumptions implicit in the simple core definition of evaluation-"Compare accomplishments with stated objectives"-may be useful here. This analysis is indebted to that of Weckworth (1969) in which he notes five implicit assumptions beneath the definition:

1. objectives are stated
2. in measurable terms;
3. accomplishments are demonstrable

4. in the same measurable terms as the objectives;
5. one knows what "compare" means, i.e., what is to be done on the basis of the findings.

A quick look at the deficiencies of these assumptions will lay the groundwork for the presentation to follow.

1) Lack of clear and/or consistent objectives in operating health programs and in new federal social legislation is the rule rather than the exception. Even if legislative intent is clear and internally consistent (an extremely difficult condition for broad-aim social programs), such intent is expressed in terms of long-term goals instead of specific, measurable objectives. There is also the problem of sorting out those public goals not meant to be accomplished.

2) Two difficulties exist in measurement: 1) interrelated objectives often conflict in such a way that achievement of one results in diminution of progress toward the other; and 2) methods adequate to measure health status and social value objectives are not always available.

3) The complexity, tediousness, or cost of appropriate measures may make demonstration of national or large scale program accomplishments counterproductive or impossible. See, for example, Sullivan's (1966) discussion of the vast practical difficulties encountered in translating some logically valid and reasonable measures of health status developed at an experimental level into a universal health index capable of being generated from data collected on a national basis.

4) What we can measure may not be translatable into progress toward an objective. What we can usually measure is some aspect of health care and what we are usually concerned about is individual or population health status. Discrete improvements in health care do not always translate

simply and directly into improvements in health status either for the individual or for population groups. In addition, there is a problem of weighing the importance of measurable vs. unmeasurable objectives or outcomes as critical overall program effectiveness.

5) Most importantly, what do any findings eventually obtained mean? How do we move from program evaluation to decisions about programs? For example, if a Maternal and Infant Care project does not meet an objective of having all mothers seen by a specialist before the third trimester of pregnancy, should it be penalized or should it receive increased resources to extend its outreach activities? If a neighborhood health center reports an increase in prevalence of chronic disease, is this a good sign or a bad one? Health conditions could be deteriorating; or people could be surviving longer and hence becoming susceptible to degenerative diseases; or the project may have developed the first accurate method of data collection and case counting in the community. If programs are doing well, will more resources increase effectiveness or will they lead to negative rates of return? And to push the issue even further, could the resources be even more productive in a non-health program? For example, even the most successful lead poisoning detection and treatment program may not be as valuable as a national housing policy or anti-lead air pollution laws.

In the next section, some of the preconditions which can either stimulate or limit evaluation research will be explored.

Part 2. The Infrastructure of Evaluation

There is a striking difference between the high quality and

sophistication of the theoretical papers in this field [health evaluation] and the lower quality research emerging from actual studies of programs. Many of the evaluations reported in the literature suffer, unfortunately, from a variety of shortcomings which reflect either lack of awareness of relevant conceptual and methodological principles or an inability to apply them properly. (Shulberg, Sheldon & Baker, 1969, p. 20)

Having established a broad concept of what evaluation can mean, I begin consideration of the frameworks within which evaluation studies are or could be designed. The details of those frameworks and of the specific content of those studies in health care will be the subject of the next chapter. But prior to any individual design are a group of conditions which may account for the frequent failure of practice so accurately depicted by Schulberg, Sheldon & Baker.

One of the most common shortcomings is insufficient attention to the assumptions underlying evaluation designs and the demands which those assumptions place on the context in which evaluation can be successfully carried out. We have already examined definitional assumptions as to the value-free or value-laden nature of evaluation and of the goal-attainment or systems model of the programs under study. In this section I want to examine a set of contextual assumptions. In development economics, the term "infrastructure" is often used to denote those underlying conditions, not part of the economic model itself, which enable the process of modernization and industrialization to take place. The infrastructure (also called social overhead capital) includes such things as power sources and water supplies (public utilities), transportation, communications systems, and education. Likewise, there is an infrastructure of evaluation, a set of pre-existing conditions and social and technical capacities which permit evaluation studies to

"take-off" into effective policy tools.

The following five elements, I would argue, are the prerequisites for effective evaluation (be it "evaluation", "evaluation research", "program evaluation", goal-oriented, or systems-oriented):

- definition of health
- health status index or indicators
- determination of health needs
- professional standards of care
- facilitating social/economic/political climate

There are two dimensions, of equal weight, to each of these elements: a cultural one and a technological one. By cultural we mean that health, health needs, patterns of care, notions of medical practice, and accepted methods of evaluation are all culturally-derived values; their use as independent variables, program objectives, or scientific techniques cannot escape being a value-laden and culture-bound process. By technological, we mean the knowledge and capacity of the system to generate data in quantifiable or evaluative form--in other words to make the cultural infrastructure accessible to the methods of evaluation research.

These dimensions are interrelated; technology is itself a part of the cultural web. But one or the other may dominate in a certain situation. Some examples may help to clarify this interdependence.

- Advances in health care and social conditions have caused morbidity to replace crude mortality as a preferred index of health status, either of an individual or the country. Technical difficulties in operationally defining morbidity (which is conceptually not the clear entity that death is) have slowed development of equally useful new national health index measures. In this case there is no cultural barrier to such development, only a technological and conceptual one. (c.f. HEW, Towards a Social Report, 1969)

- Attempts to implement water fluoridation, as a powerful dental health measure to reduce cavities in children's teeth, have been frequently thwarted because the socio-political system in some communities has not reached a facilitating consensus on the distinction between public health/public welfare and Communist plots/personal liberty. In this case, the technologies both to implement the program and to develop extremely clear and convincing evaluations exists, but the cultural environment is not always ready to accept them.
- According to research methodology, the best evaluations should be controlled experiments or quasi-experiments in design. However, ethical prohibitions against deliberately withholding a technique or drug of known or highly probable efficacy in order to create the necessary control group make such designs highly political and easily attacked. Similarly, the political process into which evaluation feeds may not have the patience to wait the years often required to do a thorough evaluation. In these cases, two subcultures (political and research) are at odds.

Let us look now at each element of infrastructure in detail and attempt to analyze their cultural and technical components.

Definition of Health

The World Health Organization definition is most frequently cited in the current literature and legislative debates:

Health is a state of complete physical, mental and social well being and not merely the absence of disease and illness.

This definition is probably without serious opposition anywhere, but because of its universality, it is also not operational. It represents only a goal for health efforts.

The importance of the definition is that it provides a basic ethical base upon which to justify efforts to improve health. As one commentator on evaluation and community mental health puts it, "We begin with the value that it is good to have good mental health." (Edgerton, 1971,

p. 92) One can understand the rhetoric of "a right to health care" as an elliptical expression of belief that health is an inextricable, although variable, part of "life", which we regard as a fundamental political right. Societies in which either life itself were not highly valued, or in which health were overwhelmingly bad or beyond human intervention, might not have basic definitions of health which would in any way lend themselves to social action, or to rational evaluation (e.g., if health were entirely a supernatural gift or curse). This definition, however, implicitly reflects belief in progress, in amelioration of social ills, indeed in social utopias.

That is why we believe this is a question for cultural anthropologists and philosophers. For our purposes, we need only establish that in general our society values health and, as we become more affluent and medically competent, that the definition of health is extended to mean more than freedom from death, pestilence and starvation.

Health Status Index or Indicators

Health status indices and indicators are an attempt to operationally define and measure health. "At present, there are no practical quantitative measures of health despite numerous attempts to develop such a measure." (Thorner, 1971, p. 526) Sullivan (1966) presents an excellent review of past and present attempts at health status indices. There is quite a bit of work being done now on more sophisticated morbidity-based indicators and indices, utilizing such formulations as disability continua, activity-restriction-days, and function. (c.f. Moriyama, 1968; Fanshel and Bush, 1970; Shapiro, 1967; de Geyndt, 1970) None of these have been

widely utilized; some of the problems they cause in policy-oriented evaluation are discussed later in connection with outcome evaluations.

Underlying concern with the development of these indicators is the belief that social and environmental conditions in some way intervene in the health status and process of the individual. One thing on which all writers in this non-paradigmatic field agree is that such linkages, while intuitively appealing, are not supported by much hard data. We know little either about how the environment causes or effects disease nor how health services influence health status.

Probably no more fundamental information would facilitate the conduct of end-result studies than knowledge of the natural history of disease, the physical, social and economic consequences of disease during well-defined intervals following onset, and the role of preventive and therapeutic medical care in altering the course of disease. (Shapiro, 1967, p. 21)

It is my opinion that an overemphasis has been placed in the past on the linkage between medical services and health. Although many of these services contribute directly to health, in some cases use of the most accepted medical knowledge and techniques yield negative results. (Bailey, 1970, p. 37)

The absence of solid evidence of the existence, direction, and extent of the causal relationships between health care and the health either of individuals or of communities has a special significance for evaluation attempts in the public domain. Too often, individual evaluation projects are saddled with the burden of developing such evidence if they are expected to document outcomes related to inputs in some program-effectiveness study. Unnecessarily negative or ambiguous results are encouraged by the conservatism of techniques of evaluation research stressing statistical significance and the rejection of null hypotheses, an approach which may be "inappropriate

in the formulation of decision criteria in evaluative research... regarding innovative programs." (Caro, 1971, p. 24) Sparer and Johnson have nicely summarized the issue in their evaluation of OEO neighborhood health centers.

...While we can be hopeful that health status will be improved because of increased use of health services, ...we should prudently resist the temptation to "evaluate" program success by [this measure].

These are research questions basic to the whole field of medical care for which basic studies have yet to be initiated on a community-wide scale and for which analytic methods or models are not available. (Sparer and Johnson, 1971, p.3)

Their response to this problem is to assume the linkage until proven otherwise, based on conventional wisdom:

In the absence of research data which relates medical care inputs to health or family functioning outcomes a program assumption can be made consistent with current social values, and indeed the values expressed by practitioners of public health, medical care, and social services. (Sparer and Johnson, 1971, p.4)

In much the same way that Thorner rather abruptly assumes the existence of a model function of health:

In concept, most health and medical care programs are predicated on the assumption that the level of health is related to the level of effort expended in providing health services, although the exact nature and strength of this relationship are usually poorly established. It is possible to express the relationship using symbolic language in the following simplified model.

$$H = f(H_s)$$

That is, health (H) is a function (f) of the effort expended on health services (H_s). (Thorner, 1971, p. 525)

In terms of the "goal-attainment" model defined earlier, this is a shift from ultimate (accomplishment) objectives, such as reducing morbidity, to intermediate (technique) ones in which cause and effect

are not at issue. Schulberg, et al. (1969) agree that this could require prior basic studies demonstrating the linkage of intermediate to ultimate objectives (the studies referred to above by Sparer and Johnson), but that in reality, "most health programs could never be undertaken if assumptions were not made about the validity of intermediate objectives and,...in the absence of facts, the public expects health professionals to use expert opinion." (Schulberg, Sheldon and Baker, 1969, p. 7)

Such a shift from an attempt to measure health to an attempt to measure effectiveness in terms of pre-existing proxy indicators can be seen in this description of the evaluation design for Children and Youth projects.

We believe that health is ultimately only a judgment of whoever has the right to decide. And our society has vested that right to decide in a select group of professional disciplines. Therefore, we wanted data to document what the outcomes or progress toward outcome of health were-- as the judges said it was. Pragmatically, even the totality of all hard data measures of health and health service delivery cannot equal the summary judgment of the professional with the right to decide. (Weckworth, 1971, p. 87)

Another type of response to the dilemma of developing practical yet meaningful indicators of health is a special use of the "systems" model (Schulberg, et al., 1969; Etzioni, 1960) For example, de Geyndt (1970) argues that the health outcome approach is handicapped since non-medical variables can affect outcome in uncertain ratios with the medical ones usually measured. He then argues that "in order to overcome this methodological roadblock, a broader framework for the assessment of quality of care should be adopted." (de Geyndt, 1970, p. 33) Attention should be paid to macro-level target population or

community impact studies, in order to open up evaluation to areas not even considered when the focus is on individual micro-level analysis. This means not only evaluation which encompasses more of the relevant aspects of program outcome; it also means that when measures of individual health status may be either impossible to derive or impractical to think of performing (e.g., physical examination), probabilistic, population measures may be workable and valuable. Infant mortality rates are not used to assess the health of the deceased infants but as measures of the general health of the population group from which they came. Recent work in epidemiology has moved toward extension of those methods to chronic diseases, morbidity conditions, etc.

De Geyndt is still operating within the goal-attainment model, but proposes expanding indicators or intermediate goals to cover more of the systematic effects of health care. The benefit is that methodological designs which analyze many variables in some multivariate, factor analysis, or path analysis way may both capture more of what people intuitively mean when they talk of "good health" and also may demonstrate the interrelated dynamics of these variables. The big drawback to the technique is that we can measure "community health", "family functioning", "stress", "emotional health", "healthy environment" only through the same kinds of intermediate objectives that we do health--and we must also assume the linkages.

A more complex and authentic "systems" model avoids or rather transcends this particular difficulty since it can study the processes of goal-selection and objective-definition as social processes, with differential results in the allocation of resources and social policy.

One could show, for example, that certain types of health status indicators or intermediate objectives were associated with politically popular, comprehensive, or radical programs while others were not.

A final type of response to the dilemma of inadequate indices is predicated on operational directedness and focuses the value assumptions (i.e. between health and health care), on areas of greatest apparent need. It is exemplified by the approach of a federal program analysis group on child health in the mid-60's:

There is no universal index of good or bad health among children. Therefore, in looking at the problem of assuring needed health care, we necessarily primarily concern ourselves with some particular health problems which are highly prevalent, which are highly adverse, and which can be mitigated or even avoided given proper health care. (Wholey, undated, p. II.3)

This response is significant for several reasons. First, it recognizes the gross inequity inherent in any approach which, given the current quantity and quality of evaluation studies, advocates program decisions based solely or even primarily on the evidence of such studies. Second, it interjects the intermediate level of professional judgments and values into health care policy explicitly: it utilizes what evaluation research data are available, synthesizes and enlarges them into a universe of health programs, and then selects targets on the basis of need and potential for impact. It has a great potential for misuse, from influence by too narrow definitions of problems or solutions, and from the capriciousness of the bureaucratic and legislative decision-making processes. In particular, it would seem to provide no mechanism to systematically consider disguised or undetected health needs, especially of the disadvantaged or powerless groups.

However, we find it a valuable tool for planning child health programs not only because, as will be discussed below, some of the accountability priorities and needs in child health are better defined than most, but also because it introduces the notion of risk or uneven need. That is, it recognizes the hard reality that not even in America are all--or nearly all--people in the enviable position of requiring assessment only by increasingly sophisticated and enlarged measures of positive health. As Sullivan has stated,

For large populations as they exist today, however, disease, illness, and death are still ever-present problems. The occurrence of these events is a more compelling guide for allocation of resources than variations in positive health. (Sullivan, 1966, p. 7)

Freeman and Sherwood (1965, p. 85) use the term "accountability" (as one of three requirements for evaluation research in large-scale intervention programs) in a similar way to mean that one should look for evidence that an important target population (by nature of its size or intensity of need) exists and that the program is in fact implemented and implemented for that group.

Most program evaluation studies, then, will make some kind of assumptions about health which justify the proclaimed value of health services and will then choose intermediate objectives such as morbidity levels, function scales, or meeting needs by which to measure effectiveness.

Determining Health Needs

If we accept in general the value of health services, we next need to know the universe of needs to assess how many services we need and whether services are being targeted properly. One analysis

of the health needs of children has identified four needs currently recognized in federal programs: (Minnesota Systems Research, Inc., 1972)

- deficits in health status
- deficits in health facilities and manpower
- specific diseases and health condition needs (e.g., Crippled Children programs, Maternal and Infant Care program, Medicaid for the blind and disabled)
- access to services (i.e. income or other barriers to care leading to deficits or potential deficits)

Need is a highly value-laden concept. It is one aspect of the judgmental nature of evaluation and cannot be said to exist in any a priori or apolitical/acultural sense. To achieve some uniformity of dimensions, the group mentioned above used proxy indicators for need which represent a set of risk factors which "are associated in a probability sense with the presence of health problems:"

- (1) selected demographic measures,
- (2) measures of poverty,
- (3) infant mortality,
- (4) immunization levels,
- (5) selected disease conditions, and
- (6) nutritional indicators.

The processes by which needs are selected as important ones, processes reflecting underlying assumptions about health, are rarely transparent or explicit. For example, note the abrupt opening of the Maternal and Child Health Care Program Analysis (HEW, 1966) laying the groundwork for the cost-effectiveness analysis to follow.

This chapter introduces the notion of a "health-depressed" area, proposes objectives, and reviews national needs in

child health... Infant mortality has long been used as the best single indicator of the health status of the community. Generally speaking, infant mortality rates are highest in areas characterized by low incomes and poor housing. Health-depressed areas could probably be best defined in terms of the variables infant mortality rate, proportion of families who are in poverty, and proportion of housing that is substandard. For simplicity, we here define a health-depressed area as an area within which the infant mortality rate is high. (emphasis in original) HEW, Maternal and Child Health Programs, 1966, p. II.1)

Not only have the authors foregone the opportunity to analyze even some limited interactions among health, health care and socio-economic variables by eliminating the poverty and housing determinants as soon as they introduced them, but they also offer no evidence (or even argument) why infant mortality rates alone will be sensitive enough to target the programs they later propose in the most "accountable" (in Freeman and Sherwood's sense) fashion possible.

The six factors proposed by the Minnesota Systems group reflect a fuller picture of community or child health status. They do not, however, reflect rising concern with the health delivery system and the interface between the individual and this system. To correct this, we can add two interrelated proxy measures of the need for comprehensive care, what Weckworth has termed the "simultaneous dimensions" of

(7) continuity in the flow of services, and

(8) completeness in the spectrum of services. (Weckworth, 1971, p. 80)

And finally, to reflect growing awareness of the interaction of patients' psycho-social needs with their health needs, we can add a measure of

- (9) patient satisfaction (this indicator is suggested by de Geyndt, 1970, p. 31).

With the exception of infant mortality, which is, as mentioned, the primary determinant of need on a community-wide basis, the nine indicators listed are in rough order of the universality of their cultural and political acceptance as legitimate indicators of need. Likewise, the technical precision and validity of specific measures associated with these indicators are in rough descending order. The first five are relatively sound, given general difficulties with morbidity indices (although, since children are generally the healthiest and least disabled portion of the population, even gross measures of reduced function or disability may have a greater than normal power of discrimination). Nutrition indicators are infrequently used, both because inexpensive, quick methods of wide-scale testing are neither available nor reliable and also because the linkages between such findings (based on clinical examinations, laboratory tests, food intake observations or food intake diaries) and nutritional status have not been investigated with any great precision at the level of subclinical malnutrition. The last three indicators are dependent for their significance on some societal or professional belief in the efficacy of health services and of certain delivery modes. They are receiving increasing attention from health policy analysts and evaluators, and if developed should prove a great step forward in expanding the value and completeness of evaluation studies.

Professional Standards

The content of medical care, its quality control, and establishment of the value of specific procedures or treatments are outside the scope of this analysis. While we will not evaluate them directly, professional standards, as reflected in the content of medical care, can still be valuable as independent variables in comparing different systems of care. For example, such criteria as the number of Board-certified physicians in a system, the percentage of women receiving annual Pap smears, or rates of elective surgery such as appendectomies, have been used to demonstrate clear differences in patterns of care between prepaid group practice and usual private fee-for-service care, with the underlying assumption being that having more of things which are professionally defined as being "good things" is an indication of better quality of care and hence of better health status among users (c.f. Donabedian, 1969).

Social/Political/Economic Climate

Health is not solely determined by use of medical services:

Health and well-being result not only from the use of medical services but from the operation of many biological, psychological, social and economic factors as well. Under ordinary circumstances it is difficult to isolate the effect of medical care from among the many interacting factors that affect health. (Donabedian, 1969, p. 23)

In general, we have as little hard data on the causal relationships between these factors and health as we do on the health services/health linkage. In fact, the data are even rarer, since the strength of our cultural belief in the supremacy of medical care for dealing

with health problems has infrequently been seriously questioned and has stifled serious research in other directions. Only in the field of public health have strong non-medical linkages occasionally been shown; in fact it has been argued that virtually all of the improvement in general health status over time has been due to improved sanitation, water supplies, shelter, and general economic well-being. This view, for example, is supported by the marginal evidence of impact by many health service program changes.

In my opinion, changes in health or medical care programs, except in some instances where a major technological or organizational change is introduced, are unlikely to produce effects on health which are easily measurable... The effects of the other variables in the model [educational level, economic level, physical environment, political situation, and others] and measurement errors are usually stronger or as strong as the effects produced by the change in the health or medical care program, and the difficulty of controlling these variables and errors may vitiate any attempt to measure the program's effects, whatever method is used. (Thorner, 1971, p. 529)

Rene Dubos has put this issue in historical perspective and addressed it more eloquently than anyone else. It is worth quoting at length from his book Mirage of Health to gain some perspective on our emphasis on medical care and scientific knowledge in our evaluations of effectiveness in improving our health and the health of our children.

The most effective techniques to avoid disease came out of the attempts to correct by social measures the injustices and the ugliness brought about by industrialization.... This achievement cannot be credited to the type of laboratory science with which we are familiar today. Rather, it was the expression of an attitude which is almost completely foreign to the modern laboratory scientist. The nineteenth-century reformers naively but firmly believed that, since disease always accompanied the want, dirt, pollution, and ugliness so common in the industrial world, health could be restored simply by bringing back to the multitudes pure air, pure water, pure food, and pleasant surroundings--the qualities of life in direct contact with nature....

In reality,...the monstrous specter of infection had become but an enfeebled shadow of its former self by the time serums, vaccines, and drugs became available to combat microbes. Indeed, many of the most terrifying microbial diseases--leprosy, plague, typhus, and the sweating sickness, for example--had all but disappeared from Europe long before the advent of the germ theory. Similarly, the general state of nutrition began to improve and the size of children in the labor classes to increase even before 1900 in most of Europe and North America. The change became noticeable long before calories, balanced diets, and vitamins had become the pride of nutrition experts, the obsession of mothers, and a source of large revenues to the manufacturers of colored packages for advertised food products....

Their [the humanitarian and social reformers of the 19th century] romantic doctrine that nature is holy and healthful was scientifically naive but proved highly effective in dealing with the most important health problems of their age. When the tide is receding from the beach it is easy to have the illusion that one can empty the ocean by removing water with a pail. The tide of infectious and nutritional diseases was rapidly receding when the laboratory scientist moved into action at the end of the past century.

The great increase in over-all expectancy of life during the past hundred years in the Western world is properly quoted as objective evidence of improvement in the general health condition. It is often overlooked, however, that this increase has been due not so much to better health in the adult years of life as to the spectacular decrease in infant mortality. The control of childhood diseases, in turn, resulted more from better nutrition and sanitary practices than from the introduction of new drugs. It is remarkable, in contrast, that little practical progress has been made toward controlling the diseases that were not dealt with by the nineteenth century reformers...The Sanitary Revolution...has had no counterpart in dealing with ailments of the adult years and of old age. (Dubos, 1959, pp. 28-31)

In view of our concern with potential conflicts between scientifically rational planning and moral or humanitarian values, it is important to note the differences in emphasis between Thorner and Donabedian and Dubos. While Dubos' open framework for attributing causal effects and for pondering questions without conclusive evidence enables him to entertain both medical and non-medical interventions, Thorner seems to see non-medical variables as non-controllable

annoyances, vitiating the continued attempt to measure the effects of health programs. Donabedian too sees his primary goal as isolating the medical care effects. Science is difficult to put second.

To return to the pragmatic world, and to summarize: social, political and economic factors influence evaluation at three levels. First, since they are among the determinants of health status and health behavior (defining illness, seeking care, following treatment, etc.), but in generally unspecified ways, they need to be considered in evaluating the impact and process of health care delivery. Second, they act as external and internal constraints on program viability; here the political realm on which Dubos is silent becomes critical. A very effective program in health terms may not maintain its relative funding level if it does not have or build a block of influential support in Congress, HEW, or the public. For example, the development, expansion, and power of the National Institutes of Health had as much to do with the political alliance of Mary Lasker, Lister Hill and John Fogarty as it did with the usefulness of such organizations in doing socially valuable basic medical research. (Drew, 1967)

The third and final level considers this context as an alternative to direct provision of health services for intervention to improve health status. Consideration of even a few of the possible trade-offs between intervention modes is impossible here, but in the long-run full examination of such trade-offs and alternative explanatory theories is absolutely essential. As one example, it seems perfectly plausible that we could do the most to improve children's cognitive skills and achievement levels by investing money now put in public education in universal prenatal and infant care, and assuring good nutrition for

everyone; and then using schools as screening and treatment centers. Or perhaps putting it into a guaranteed annual income for all Americans. Technical capacity to actually make such trade-offs on the basis of cost/benefit analysis or other "scientific" criteria does not yet exist. (c.f. Rivlin, 1971, pp. 51-63)

Chapter Three.
The Evaluation Process:
Existing Frameworks and a Tentative Model for Health Services Evaluation

Into the dynamic cultural milieu, whose important variables we have described under the term "infrastructure," the undaunted planner interjects the structure and process of evaluation; be it evaluation research, program evaluation, or just plain evaluation. Abstracting from the multiple, and often conflicting, definitions discussed in the last chapter, I would like to examine the substance of evaluation as it gets practiced--or rather as it is proposed to be practiced.

As in the section on infrastructure, I will try to illuminate some of the differences between goal-attainment and systems models of evaluation. Also of concern will be the ways in which value judgments are assumed, incorporated or ignored in various types of evaluation and the openness of those types to qualitative issues.

Agreement on the essential parts of a good evaluation or on the words used to describe different types of evaluation is no greater than it is on basic definitions. This has presented some difficulties in organizing this section. I did not wish to merely describe each author's scheme separately, both because they are not sufficiently different conceptually, and also because I wanted to attempt to synthesize them into a comprehensive model of evaluation which would at least theoretically bridge the gaps between rigorous scientific experimental designs (i.e. true evaluative research) and designs focused on processes of policy making, value judgments, and so on. But, since much of the

literature is abstract and does not give concrete examples of the types of evaluation described, it is very hard to be sure that two people are in fact talking about the same thing. Given the relatively low intellectual quality of many of the pieces, particularly those specifically dealing with health program evaluation, it would be an arid exercise to rigorously catalogue the various schema. What I have done, therefore, is to begin with brief explications of the major formulations I have found to indicate the types of vocabulary and levels of analysis used. Then I have gone directly to the framework of six components that I believe essential to an ideal evaluation of child health programs. In discussing these components at some length, I refer to the previously mentioned constructs as they seem to be either synonymous, parallel, or contradictory.

There was one further difficulty in combining the various studies. There are large gaps in terminology, methodology, and in level of analysis among three groups of authors:

- 1) people who consider themselves evaluation researchers or program evaluators. They usually do non-federal studies or models, focused on institutional reform and social change but from the vantage point of one program (e.g., venereal disease education; anti-smoking campaign; new type of health manpower) or one discipline (e.g., community mental health; hospital operations; prepaid medical care). The published studies tend to be heavily concerned with methodology, use of experimental, quasi-experimental, or field trial designs and with the problems of evaluator/program operator interactions and the ways in which evaluation studies can be best and most widely utilized.

- 2) People who are basically policy analysts focused on program

evaluation as one level of the governmental decision-making process. Frequently although not always from the PPBS school, the authors are concerned with interprogram trade-offs (e.g., Headstart vs. Homestart; family planning versus intensive newborn care). They tend to discuss not research methodology, which they often seem either to be ignorant of or to assume to be adequate to provide raw evaluation data of high quality, but rather with cost-effectiveness, cost-benefit and efficiency studies, the role of evaluation in policy making, and so on.

3) People who are writing directly about health programs and evaluations. This work may be at either of those foci mentioned above; there seem to be few health studies, other than in mental health or health education. Given their greater specificity, health evaluation articles tend to give more details on the content of the evaluation, but in doing so, they often seem to forget some wider issues. In addition, many of these studies have been done at levels of analysis which we are omitting here, namely of the content of medical care, i.e. the efficacy of specific treatments or procedures in curing or alleviating specific disease conditions. This type of research resembles basic biomedical research and generally involves greater attention to experimental design or to clinical study procedures than do the broader-based studies in which we are interested. I mention this because the picture I paint of the state of the art of evaluation may seem overly dismal if one does not consider the underlying research which may someday influence program action by finding cures or better treatment methods.

Some Frameworks of Evaluation

In this section, I will try to give some indication of what sorts of things people study when they do evaluation, what categories of knowledge they consider useful or necessary to the process. I particularly want to present some of the fuller schemes for evaluation, since they are not identical with the one which I will propose and thus I will need to refer back to compare their terminology and categories with mine. Writings which deal with only one or two types of evaluation can be more easily and understandably incorporated into that later discussion and will not be referred to here.

Schulberg, Sheldon and Baker (1969), in their reader on health evaluation studies, suggest that evaluation research can be performed at several levels:

1. evaluation of effort: comparing a local or individual program against local or national standards
2. evaluation of performance: documenting outcomes of program; assumes services correctly delivered
3. adequacy of performance: to what extent was total problem solved
4. evaluation of efficiency: can identical result be achieved at lower cost (Schulberg, Sheldon & Baker, 1969, p. 8)

In all these cases, the underlying model is the goal-attainment one of comparing performance with stated objectives. This model "presents an easily conceived relationship between a specific service and an ultimate effect, while the methodology to be used in assessing outcome generally is drawn from the well-established designs of researchers." (Ibid., p. 9)

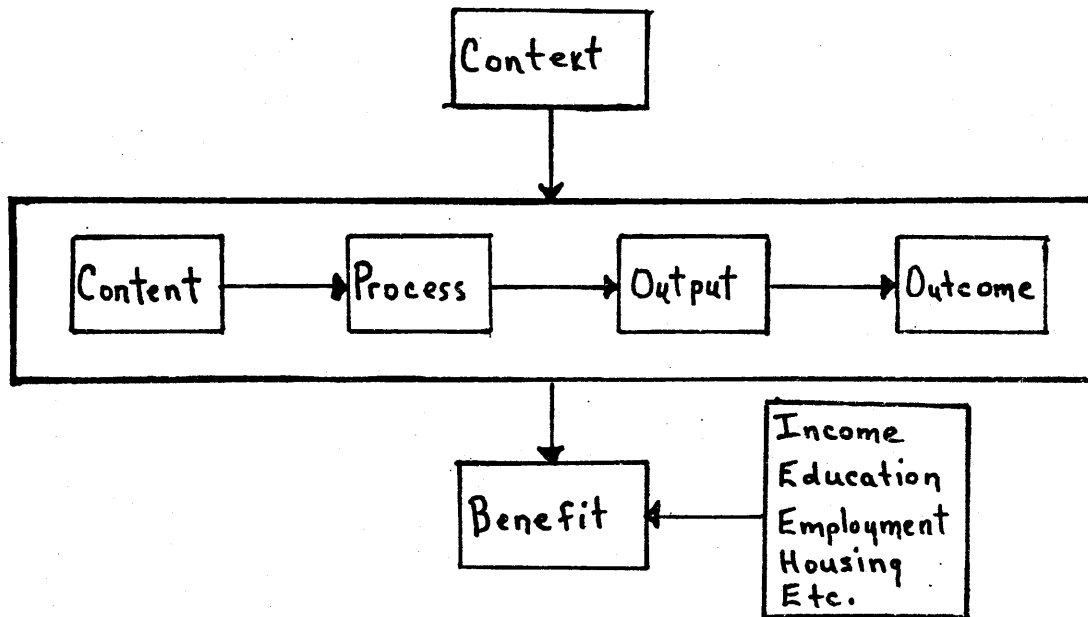
Wholey et al. (1970) discuss four types of evaluation pertinent to the analysis of federal social programs:

- 1) Program impact evaluation: assessment of overall effectiveness of a national program or the relative effectiveness of two or more programs aimed at the same goal. Defining and measuring output variables and comparison groups are critical.
- 2) Program strategy evaluation: Assessment of relative effectiveness of different techniques used in a national program. Defining and measuring appropriate environmental, input, process, and output variables, as well as suitable analytic models, are critical.
- 3) Project evaluation: assessment of effectiveness of individual project in achieving objectives. Like program impact studies, output variables are measured against either comparison groups or against baseline conditions.
- 4) Project rating: assessment of relative effectiveness of different local projects in achieving program objectives. Defining and measuring environmental variables and quick, short-term output measures are usually used.

This analysis done at the federal level does not stress methodologic issues, but clearly makes large assumptions about the ability of subsidiary systems and institutions to generate data and to provide operational analytic models. The authors, in contrast to Schulberg et al., do not consider cost analysis as a part of evaluation itself. They differentiate it, along with other evaluation-related activities, because it is concerned with program inputs, not outputs. They do view it as essential to good evaluation but do not make clear at what steps in the process it should be the critical factor in decisions about program effectiveness. (Wholey et al., 1970, p. 27)

In one of the most fully developed programmatic proposals for health program evaluation, de Geyndt (1969) and Weckworth (1971) present a framework which is specifically "descriptive" conceptually. The Systems Development Project (now called Minnesota Systems Research, Inc.) which Weckworth directs and de Geyndt works with makes elaborate distinctions between prescriptive models which involve value

conflicts (and are hence inappropriate for studying the effectiveness of a total health care program) and descriptive, explanatory, and predictive ones. (de Geyndt, 1969). The model they developed is, they believe, basically a descriptive one with some explanatory powers, fewer predictive ones, and no prescriptive intentions. It owes much to organizational theory and can be represented graphically as follows:



De Geyndt goes into great detail on the types of variables one could analyze in each of these categories. In later sections, some of these will be listed; at this point, only his basic definitions of the categories are presented:

- 1) Context: the setting; the internal and external environment in which a project functions. Includes political, economic, and social dimensions of the environment. Context is important because it acts to constrain the program's achievement of objectives, to limit its effects, and to bound its rationality.
- 2) Content: program elements provided, services rendered, or activities carried out by the project.
- 3) Process: describes the actual functioning and operation of project. Represents the transformation of the inputs of

context and content into outputs. Or, "process documents the throughputs, or the combination of the different flows and their relationships." Differentiates the "anatomy" of process (static elements) from the "physiology" of process (the interrelationships and degrees of interconnectedness).

- 4) Output: the result of the performance of an activity (e.g., a filled tooth, a social worker's visit, an X-ray, etc.) Distinguishes intermediate and final outputs, then notes that the former are actually inputs (e.g., all of the above examples except perhaps the filled tooth).
- 5) Outcome: attainment of a level of health by the patient. Can be the result of one output, of a logical sequence of outputs, or a combination of widely different outputs. A time horizon needs to be introduced since one measures health along a time continuum (i.e. a negative chest X-ray would be considered a current good health outcome, with no further process of care needed; but if it were not repeated yearly or so, then in five years the outcome meaning of the test would no longer be clear). Individual health outcomes are distinguished from non-individual ones, such as the spin-off effects of the program itself on the community, on health education, on pay scales for health workers, etc.
- 6) Benefit: interrelates the effects of provision of health services with other aspects of the total environment towards achievement of a state of positive well-being. This enterprise is closely related to the development of social indicators.

This framework, which was proposed to study the federal Children and Youth Program, is both confused and insightful. De Geyndt notes that the methodology would not incorporate any attempt to measure benefit. That methodology, by the way, is heavily reliant on data-collection techniques and multivariant and correlational analyses. The framework suggests the difficulties one encounters when one tries to encompass everything under one evaluative roof. The notion of "output" is particularly unclear. Part of the problem is that to achieve clarity within your own model, you must violate traditional usage in so many ways that no one may understand what you mean. As de Geyndt notes about the ambiguous meaning of "intermediate outputs": "In the health

field we have a long standing tradition of using inputs as outputs and documenting our activities by counting inputs." (p. 11)

In a later article, de Geyndt proposes a similar framework, this time a more general discussion of approaches for assessing the quality of care (de Geyndt, 1970). First, he notes that care can be analyzed at three, increasingly comprehensive and complex levels: patient care (episodic disease care), medical care (preventive medicine and episodic disease care), and health care (comprehensive care, quality of life). Then he discusses five approaches for assessing the quality of care, which may be more or less appropriate or useful at those different levels of analysis:

- 1) Assessment of Content: usual meaning of "quality of care"-- is medicine practiced properly. Relies on quality control committees (in hospitals or group practice), medical audits, tissue review, etc. Quality is defined as the degree of conformity with preset standards and focuses almost entirely on patient care.
- 2) Assessment of Process: evaluates the total management of the patient (preventive, diagnostic, curative, rehabilitative). Stress on coordination of care and care-givers, logical sequence of elements of care.
- 3) Assessment of Structure: analysis of organization of facilities and equipment and manpower.
- 4) Assessment of Outcome: attempt to find measurable aspect of health status of individual or group and to determine change in this status as result of content, process, and structure of care. (corresponds to earlier category "outcome")
- 5) Impact: measurement of the impact of health care and health services system on the health of the community or the nation. (Roughly corresponds to earlier category of "benefit" and to analytic level of "health care")

In this second formulation, de Geyndt has omitted "context" as a level of evaluation and it is only implied. This is in keeping with my formulation of the infrastructure of evaluation, since de Geyndt's term

was analagous to my element of socio/economic/political context. In this second piece, what he first called the "anatomy" of process is now termed "structure" and the ambiguous term "output" has been subsumed under both "process" and "outcome."

Freeman and Sherwood (1965) define evaluation as the process of using findings from social action experiments as the basis of efficient resource allocation. They note three requirements for evaluation research in order to perform this function.

1) Determination of efficacy. For this purpose, they stress the need for experimental or quasi-experimental design, direct behavioral measures (instead of more easily assessed attitude measures) for the impact model, randomization of study groups when feasible (more often than many suggest), etc.

2) Accountability. This refers to the determination that there is in fact "a target population that can be dealt with by means of a program; that this population is important either because of its size or the intensity of pathology; and that the project program for the target population actually is undertaken with them." (Freeman & Sherwood, 1965; reprinted in Schulberg, et al., 1969, p. 85)

3) Determination of efficiency. "In terms of all programs, the efficient one is that which yields the greatest per unit change, not the one that can be run at the least cost per recipient." (Ibid., p. 86) Efficacy here is roughly synonymous with "outcome" (de Geyndt) or "performance" (Schulberg et al.) while accountability incorporates both measures of "process" (are people actually being served as the plan proposed) and of "impact" (what effect does the program have on the

total health of the community). Efficiency is usually thought of as identical with cost analysis (Wholey) or cost-effectiveness (Schulberg), but Freeman and Sherman are the only ones who specifically state that their criterion for efficiency is cost per unit change rather than cost per unit treatment or cost per person eligible. This is a critical distinction which should be made explicit more often (the measure one chooses being clearly a matter of value judgment).

Finally, let us look at the criteria for acceptable evaluation which Stufflebeam et al. (1970) propose. Their definition of evaluation is a broad one, viewing it as a process of gathering and organizing information useful for judging decision alternatives. This information needs to have representational "goodness," practical "utility," and to be prudent. (Stufflebeam et al., 1970, p. 26) Goodness here refers to scientific criteria.

1. Scientific Criteria

- a) internal validity
- b) external validity
- c) reliability (consistency of information)
- d) objectivity (publicness of information)

2. Practical Criteria

- a) relevance
- b) importance
- c) scope
- d) credibility (a function of trust or belief in evaluator)
- e) timeliness
- f) pervasiveness (of findings)

3. Prudential Criterion

- a) efficiency (i.e. cost-effectiveness)

The scientific criteria are, interestingly enough, like those generally thought of in connection with evaluation research design (c.f. Campbell

and Stanley, 1963, for the authoritative discussion of requirements for internal and external validity). What Stufflebeam et al. seem to suggest, then, is that this rigorous design must be coupled with broader, non-scientific criteria. "Importance" and "scope" are similar to "impact" (de Geyndt) and "accountability" (Freeman and Sherman). The issues of relevance, credibility, and timeliness are discussed by most writers although not directly incorporated into their conceptual frameworks. For example, de Geyndt (1969) concludes with a paraphrasing of a paradox expressed earlier by Marris and Rein (1967, p. 3):

"The evaluation...becomes more specific and more accurate the longer we wait, but this waiting period and the search for accuracy conflicts with the urgency of decision-making. Therein lies the dilemma of the social researcher: to say something to the point, before it is too late, and the need to make sure that what he says is true." (de Geyndt, 1969, p. 16)

A Tentative Model of Health Services Evaluation

None of the constructs presented above incorporate all of the elements necessary to do evaluation in such a way that scientific reliability, ethical integrity, and policy-making practicality are equally and fully respected. I do not pretend to have that model. What follows is a first attempt: a pulling together of the content of and the gaps in present models in order to suggest the outlines of the types of information, methodology, and analytic concepts which I believe should inform good evaluation.

There are six components to this tentative model. I will discuss each one at some length, documenting the ways in which this type of evaluation has been tried, the aspects of health care and health services delivery for which it is most useful or essential as an

analytical tool, and some estimates of the potential for further development and refinement of the techniques requisite for it.

There is a seventh aspect of health care which, while absolutely essential, seems to me to lie outside the purview of evaluation research or program research. This is the content of care at a medical or biological level. First, the content of care is in many ways part of the infrastructure of evaluation. The level of biomedical knowledge, the technological ability to build heart/lung machines or complicated monitoring systems for intensive care, and the cultural beliefs which put faith in vaccines, tranquilizers, and surgery, but not in acupuncture, herbal remedies, or witch doctors--all of these are usually taken as givens in assessing the performance of a health program at any one point in time. Certainly these levels of knowledge, degrees of technical competence, and cultural beliefs change over time, sometimes quite dramatically and quickly. But the research and activity which lead to these changes usually take place in settings other than social action programs with direct service objectives. They are evaluated by clinical trials, biomedical research (either basic or applied), or by the weight of public opinion. In this sense, the content of medical care is reflected in the professional standards (including not only those of the AMA but other more liberal or iconoclastic values) of the infrastructure. Evaluation of the extent to which an on-going program meets these standards is an important part of program evaluation and could be considered a subcategory of process evaluation. There are special evaluation tools for such process studies, such as peer review committees, medical audits, record reviews, etc. These are beyond the scope of the

discussion here, primarily because critically analyzing the methodology of such studies presumes more detailed medical knowledge than I possess. Some examples of this type of evaluation are the work of Morehead et al. on the neighborhood health centers (Morehead, Donaldson, & Seravalli, 1971) and of Donabedian (1966, 1968). In a strict sense, "quality of care" evaluation begins, and in limited studies ends, with this type of analysis of adherence to professional standards.

The following six components, then, seem essential to a complete evaluation of health programs beyond the level of content of care:

--Organization

--Process

--Outcome

--Impact

--Advocacy

--Actionability

The first four components are those most commonly found in evaluation designs; the last two reflect the criteria most frequently lacking or considered only peripherally. Outcome (or "end-result" or "accomplishment") and process measures predominate (c.f. Shapiro, 1967; Donabedian, 1968; Strauss & Sparer, 1970; Freeman & Sherwood, 1965; Fanshel & Bush, 1970), while some designs incorporate three or four (Weckworth, 1971; De Geyndt, 1969, 1970; Schulberg, Sheldon & Baker, 1969; Wholey et al., 1970).

Organization

Evaluation of organizational variables is similar to de Geyndt's term "the anatomy of process," which includes the static elements of 1) specialization or division of labor, 2) standardization (of treatment

patterns), 3) formalization (of criteria and standards), and 4) centralization (degree of delegation of authority). (de Geyndt, 1969, p. 10) In general, organizational variables include facilities and equipment, manpower and staffing patterns, organizational arrangements (formal and informal), and financial mechanisms (c.f. de Geyndt, 1970).

Organizational or structural analysis is a rather encapsulated branch of program evaluation with stronger links to sociology and to organizational theory than to health care research. The methodologies used are also generally better developed and more uniform than most types of health care evaluation, and thus the evaluation process is more likely to be comprehensive, comparable to other studies, and accepted by outside critics (c.f. Stufflebeam's criteria of "credibility"). This type of evaluation has been particularly popular in the study of hospital operations (to be expected since the focus is on an institution rather than a type of service or a group of people) and in the mental health field. In the latter area, this reflects a traditional emphasis in mental health on hospitalization and professional care as modes of treatment; hence elaborate studies of the effects of ward management, staff relationships, intake procedures, or use of different types of caregivers are frequent in the field. I will not go into detail in either of these areas however, as they extend too far afield from my concern with direct service programs. Hospital studies tend to focus around institutional goals (growth, efficiency, modernization, solvency) rather than patient goals, and I have reluctantly excluded mental health programs in general because of time and space limitations.

The methodologies used in these areas, however, can be applied to

different settings with beneficial results. For example, in their attempts to devise a conceptual framework for the evaluation of the Children and Youth Program, Weckworth and his colleagues (de Geyndt, 1969, pp. 3-6) found that studies of business firms and institutional operations provided them with the most useful insights. Organizational analysis has much in common with a systems orientation to evaluation in general. Etzioni, as we noted earlier, states that the starting point of evaluation should be a "working model of the social unit which is capable of achieving a goal." (Etzioni, 1960). Etzioni's ideas are rarely integrated into health program evaluation. This may be due in part to the general lack of intellectual depth of many writers in the field of evaluation, who tend to stress one perspective at the expense (and often apparently in ignorance) of others. But it may also be due to more pragmatic concerns. Organizational analysis is a rich field, but its techniques are not particularly well designed to demonstrate the causal linkages between organizational patterns and health status. Certainly any outcome or process study of necessity tests certain organizational patterns; but it is not necessary to explicate the full dynamics of a system each time one studies an outcome. Or is it? One thing which should be done, and which unfortunately is frequently omitted, is an explicit description of organizational patterns under which other types of evaluation are done, so that later critics can make their own value judgments as to whether any of those factors might be the real causes of demonstrated effectiveness or negative results. For example, studies of the health and health behavior of the poor in this country all too often draw conclusions from utilization rates or amounts of

money spent on health care or make comparisons with British utilization data without acknowledging that under our system of health care access is not guaranteed. Hence utilization is a function of income, geography, or other barriers as well as of the individual's attitudes or health.

There are few studies which directly link organizational factors to the special needs of children. One area in which work has been done is the effects of hospitalization on young children. Although this work grew out of psychological and psychiatric interest in the severe damage which often was done to children's psychological functioning (c.f. the term "hospitalism"), the results of the work have had major impact on practice. Pediatric hospitalization rates have been declining for years and current hospital planning stresses shortness of stay, living-in arrangements for parents, day-hospitalization, and other such procedures. (c.f. Spitz, 1945, 1946; Shaeffer & Callender, 1959; Dimock, 1960; and Bergman, 1965, for discussions of the psychological bases for the concern, and Shore (Ed), 1967, for one example of how this knowledge is influencing planning decisions about pediatric hospitalization). Broader studies which examine the ways in which the national health care delivery system impacts on children will be discussed in a later chapter.

Process

Process evaluation examines the technical aspects of quality of care (content) in a sequential context of interaction between a patient and the delivery of health services.

The process view tends towards the concept of the "whole patient" and evaluates not only the work of the physician but also the contributions of other health workers. Total management of the patient encompasses the prevention of illness, the arrest of

pathological processes, and physical rehabilitation... Coordination of the process becomes important and the enabling role of administration is crucial... Patient care is supplemented by medical care, and the term "delivery" is a decisive part of the quality for it is the system which enables the process to take place. (de Geyndt, 1970, p. 26)

Process evaluation is analogous to what Wholey et al. (1970) call program strategy evaluation and what Schulberg et al. (1969) term evaluation of "effort." Suchman appears to suggest that process analysis is the addition to strict "evaluation research" which produces program evaluation (c.f. Wholey et al., pp. 15 & 94), by providing enough insight into which components of the program are causing the results measured so that policy decisions can be made.

Existing criteria for process evaluation are not the complete, coordinated set suggested by de Geyndt's rather utopian model of the care process. In a descriptive sense, there are the elements which de Geyndt (1969) classifies under "content":

- program elements covered and their priority (preventive, diagnostic, curative, ameliorative, restorative, rehabilitative, and emergency care)
- to whom provided (eligibility criteria, catchment area)
- by whom rendered (types of personnel)
- formal and informal linkages to other health or social services providers

For young children, for example, one would expect to see priority placed on preventive and diagnostic services either available to all children or targeted to high risk groups, provided by pediatricians or other specially trained personnel, in ambulatory settings, with close ties to schools, social services, etc.

Second, there are the dynamic elements of process, what de Geyndt

terms the "physiology" of process, "the interrelationships or the degree of interconnectedness of the variables and attributes subsumed under context, content, and the anatomy [static] part of process."

(de Geyndt, 1969, p. 10) These interrelationships include:

- staff referral mechanisms
- interchangeability of skills
- innovations in delegation of tasks (e.g., pediatric nurse practitioner, community health aides or outreach workers)
- actual patterns of treatment ("who does what, when, and to whom")
- decision-making process.

Process dynamics can also take a patient-oriented approach and focus on the progress of the patient through the system, and on the elements in the system as they affect the patient. The reporting system developed by the Systems Development Project for the Children and Youth program (and now being extended to the Maternal and Infant Care program) is a good example of this type of evaluation. The system tracks registrant cohorts from intake to health assessment, any ordered treatments, and to health supervision (long-term care management, routine preventive services) in several functional areas (medical, dental, nursing, psychology, nutrition, social service, speech and hearing, physical therapy, and occupational therapy). While requiring the generation of substantial quarterly data by the individual projects, it appears to be a very sensitive instrument to track the actual progress (and regressions or returns for care) of the patients and of registered persons who require no immediate health care. (Weckworth, 1971)

More selectively, some process evaluation studies have used specific

procedures in the process of care delivery as proxy measures for the effectiveness of that process. The occurrence of such a procedure is taken as an indication of a certain amount of health care, and hence an improvement in health status. This approach has appeal because it greatly simplifies data collection and reporting systems (compared to, say, the complexity and extent of the data required to support Weckworth's model). On the other hand, such simplicity may be achieved at the expense of validity. Typical process "output" measures include visits per years (broken down into types), immunizations, diagnostic screenings, dental visits, number of registrants, hospitalization rates and length of stay, etc. The measures are basically equivalent to utilization measures. But process outputs have to be related to need and to medical practice standards to be meaningful. For example, a system can generate a large number of visits per year per registrant because initial screening and assessment is inadequate and professional care rushed and symptomatic only; hence registrants use the facility frequently, as an emergency room. Or visits may be the result of excellent initial screening, followed by an extended program of rehabilitation services, health education, and community outreach. Without data on the content of process itself, crude utilization data is virtually impossible to interpret correctly. Only for very specific, completely standardized procedures such as vaccinations can we be sure that different programs or systems mean the same thing by the same term.

Utilization data are fairly easy to assemble and compare but very difficult to interpret. In the first instance, utilization of services depends on the occurrence of a condition which must be recognized by the patient, or by those responsible for him, to require medical attention. Once care is initiated by the patient, further use of service is largely determined by the manner in

which the physician manages the case, subject, of course, to continuing cooperation by the patient. (Donabedian, 1969, p. 10)

Unfortunately, few analysts are even slowed by this caveat. In particular, use of utilization data to determine underlying need (e.g., are the poor sicker; do they need education to use services in a more appropriate way) will generally both underestimate need and misinterpret behavior if not controlled for process variables--especially eligibility, place and time of care provision, and program linkages.

There is professional disagreement as to the value of process evaluation per se. Donabedian has stressed its vital function for the feedback of evaluation into program change and planning.

Outcomes...provide the most convenient and valid indicator of whether the allocation of resources under one system of care (or in the practice of a given provider) is more or less effective or efficient. But more detailed analysis of the medical care process itself is necessary if the reasons for lack of effectiveness or low efficiency are to be pinpointed and corrected. (Donabedian, 1968, p. 184)

On the negative side, there is the continuing struggle to find dynamic process measures which are both valid and obtainable on a systematic basis.

The development of measuring instruments to express the inter-relationships between the dynamic elements of the care process and the interactions among participants represents a major difficulty. (de Geyndt, 1970, p. 27)

Wholey et al. (1972) go even further and note that

the description of the treatment to which individuals have been exposed can be complex and very expensive...Precise evaluation of the effectiveness of more complex treatments within an operating program, especially evaluation that would tell why a particular local project worked, are beyond existing evaluation methodology. (p. 99)

In the next section, a new methodology, devised by the Institute of Medicine, National Academy of Sciences (National Academy of Sciences,

1972) which incorporates both process and outcome analysis will be presented in some detail. This technique, currently being tested on evaluations of ambulatory care in neighborhood health centers, is an extremely promising, carefully thought out, and justified approach which may eventually temper the pessimism of current observers.

Outcome

Outcome evaluation is usually considered the ideal model, since it attempts

to determine the change that has taken place in...a measurable aspect of the health status of an individual or of a group of individuals...as a result of the conditions that effect the content, the process, and the structure of health care. It posits that the ultimate criterion to judge the quality of health care rendered lies in an alteration of the health of the recipient of this care. (de Geyndt, 1970, p. 29)

Outcome evaluation is what Schulberg, Sheldon & Baker (1969) call evaluation of performance, Wholey et al. (1970) term program impact evaluation (or project evaluation), and Freeman and Sherman (1965) call efficacy.

Since outcome evaluation can be seen as a programmatic attempt to validate our conventional wisdom that medical care has a non-random effect on health status (of individuals or communities), it is not surprising that this "ultimate criterion" is more easily conceived of than established.

There are problems, suggested by our discussion of infrastructure, with indicators and measures of change.

The desirability of determining quality of medical care by its effect on some measurable aspect of health is matched by the pessimism among researchers about the possibility of success in dealing with the issue. (Shapiro, 1967, p. 7)

But this is a relatively new field, almost all of the work in the area having been done in the last fifteen years (de Geyndt, 1970, p. 32), and continues to be very popular. Shapiro (1967), in a review of past and on-going studies of end-result evaluation concludes with measured optimism on progress toward the development of more adequate measures and more comprehensive end-result studies.

De Geyndt (1970) in a review of outcome studies, found five main types of measures used:

--mortality

--morbidity

--disability

--physical functioning (e.g. rehabilitation)

--social functioning

Other measures used for certain types of studies include surgical procedures* and birthweight and Apgar scores. De Geyndt also notes the important measures suggested by Donabedian of patient and provider satisfaction.

In his review and synthesis of health status indicators, Sullivan (1966) reduces the diversity of attempted morbidity measures into three categories:

--clinical evidence

--subjective evidence (e.g., health surveys which ask respondents to report illnesses and utilization behavior)

--behavioral evidence (reported evidence with more corroborative

*Rates of appendectomy and hysterectomy, while really utilization rates, because of their control by physicians are good quality measures for hospitals and some systems of medical care.

bases: loss of time from work or school, institutional confinement, activity restriction, medical expenditures).

Fanshel and Bush (1970) report some promising work in developing function, which they find from a review of the literature as "central to any generalized notion of well-being," as an operational, quantifiable indicator of changes in population health status over time. They formulate their index function as a continuum, divided into a range of mutually exclusive categories. Every member of a population group can be classified as belonging to one and only one category at any point in time; movement through these categories (positive or negative) can then be observed and measured. The extent and duration of variance from the positive pole of complete well-being becomes a measure of the suboptimality of the health status of the group or individual. This index is conceptually similar to the process assessment methodology designed by Systems Development Inc. (de Geyndt, 1969; Weckworth, 1971) for the Children and Youth Program, in which patients are charted as they move through clearly delineated parts of a health service. One could imagine a combination of these continuum scales which would attempt to directly establish the correlations between progress (or regress) through health care and levels of health.

There is one major problem with any disability or function scale which will not be solved through even such careful, mathematically precise work as that of Fanshel and Bush. This has to do with the behavioral basis of the measures and the difficulty in obtaining reliable assessments of them. For example, Fanshel and Bush's scale includes "well-being, dissatisfaction, discomfort, minor disability, major disability, disabled, confined, confined-bedridden, isolated,

coma and death"; only the last four or five would be non-behavioral in determination. As Sullivan points out:

...while a period of restricted activity is a more objective event than a report of "feeling ill" it also has a subjective aspect which must be recognized. The decision to reduce his usual activities reflects the individual's attitude toward illness and self-care, the expectations or demands of his family, his employer and his associates, his knowledge or beliefs about the symptoms present, and other social and cultural factors...Disability occurs in a social setting and, like other social phenomena, cannot be measured in isolation from the setting. It has social consequences regardless of the nonmedical factors which may influence its occurrence. A condition which disables a salaried worker may not disable a person paid on a daily basis--this makes the disability no less real for the salaried worker and his employer. Disability measures reflect the impact of morbid conditions as they influence the social participation of members of the population. In this respect they measure an aspect of morbidity important in any evaluation of the health status of a population. (Sullivan, 1966, p. 18)

These types of measures are powerful in that they reflect the human reality that the same objective health status may have different meanings for two different people. The loss of a leg would be a disabling event for an athlete or a policeman; it might be only a disability for a sedentary person. Cervical cancer which necessitates removal of the reproductive organs may have stunningly different impacts on the lives of a young woman planning to have children and a post-menopausal woman. What the measures don't equitably reflect, however, is the extent to which individual health status is not a personal, behavioral "choice" but a social one. Ability to take sick days, adequacy of health insurance, housing suitable for bedrest or recuperation (e.g., not having to walk up several flights after a heart attack or with a heart condition)--these and other variables intervene in the functioning of us all as healthy beings. None of

the work which I have seen on disability or function scales seriously deal with these biases.

Even putting aside behavioral scales, outcome evaluation depends on the infrastructure of health status indices. The general lack of operational concepts in that area has already been discussed at length. Although Shapiro (1967) also discusses methodological problems of outcome research as adequate control groups, selectivity bias, and sources of observational data, both he and de Geyndt (1970) agree that the "major deterrent has been the difficulty in defining the measures to be used." (Shapiro, 1967, p. 28)* De Geyndt also mentions a temporal aspect of the problem: measures other than mortality or acute morbidity cannot be made at the time care is rendered. Thus they cannot be used for cross-sectional comparisons of outcomes, except retrospectively in longitudinal studies. Such studies require a long "lead-time" (c.f. 20 years for the Framingham heart disease study). The costliness of this methodology, as well as the problems with relating evaluation to program planning cycles, are apparent.

To counteract some of the pessimism suggested so far, there is one promising innovative methodology which combines content, process, and outcome evaluation with which I am familiar and which I would like to describe in some detail.

Institute of Medicine Methodology for Evaluation

Under contract to HEW, the Institute of Medicine of the National Academy of Sciences started a program in 1969 called "Contrasts in

*We will discuss methodological issues separately later.

Health Status: A Comparative Inquiry into the Health Needs, Barriers and Resources of Selected Population Groups," which had three objectives:

1. To analyze differences in the health status of contrasting population groups.
2. To relate differences in health status to demographic, environmental, biological, medical care, economic, and behavioral characteristics.
3. To compare the effect of various models of health care delivery on selected population groups. (National Academy of Sciences, 1972, p. 1)

The basis of the study design by which differences in health status were to be measured was the premise that "specific morbidity conditions ...or health problems can be used as indicators of health status and outcome, which, in turn, can be related to the type and quality of care rendered." (Ibid) NAS termed these conditions or problems tracers, and developed a methodology based on them relevant to an evaluation of ambulatory health services. This involved both conceptual tasks (defining criteria for selecting tracers, selecting a set of tracers for ambulatory care, fully describing criteria for primary care associated with each tracer, and adapting the general methodology for a practice setting, specifically neighborhood health centers) and field activities to test the design. Only the conceptual tasks have been both completed and reported on; results from the completed field study should be available soon.

The process of defining tracers is worth examining in depth. The study group explicitly states that the traditional use of morbidity conditions as indicators of process or outcome variables needs to be expanded to include concurrent "assessment of provider and recipient behavior and their interaction." (p. 8)

The ultimate outcome of health care is not only a reflection of the type of care delivered by the provider, but also a mirror of the characteristics and behavior of the patient population. (Ibid.)

A formalized decision-making tree was developed containing the criteria by which tracers were chosen. Those criteria, in order of their application are:

1. The functional impact on those affected should be significant.
2. A tracer should be relatively well-defined and easy to diagnose in both field and practice settings.
3. Prevalence rates should be high enough to permit the collection of adequate data from a limited population sample.
4. The natural history of the condition should vary with utilization and effectiveness of medical care.
5. The techniques of medical management of the condition should be well defined for at least one of the following processes: prevention, diagnosis, treatment, and/or rehabilitation.
6. The effects of nonmedical factors on the tracer should be understood.(i.e. its epidemiology) (Ibid., pp. 10-13)

These criteria represent an attempt to systematically screen health knowledge to choose those areas in which we are more certain of finding need, less uncertain about the influence of nonmedical factors, more concerned about severity of need, and less hesitant about asserting the usefulness of medical intervention. This selective confidence is a value judgment of the best kind--one which pushes knowledge and scientific veracity as far as possible but which in the end does not refuse to make reasoned choices on the basis of incomplete evidence. We know too much to select morbidity conditions at random, but we know too little to use them all as indicators. I do not know enough about medicine to know how good the tracers selected are or if the choices considered were wide enough. But the report provides extensive

documentation on the etiology, epidemiology, functional impact, and medical management of the tracers selected so that one could review the case independently.*

Of the seven tracers selected, four are judged to be relevant for evaluating health care to young people:

- Middle ear infection: ages under 5 and 5-24, both sexes
- Associated hearing loss: ages 5-24, both sexes
- Visual disorders: ages 5-24, both sexes
- Iron deficiency anemia: ages under 5, 25 and over, both sexes
- Hypertension: 25 and over, both sexes
- Urinary tract infections: 25 and over (female); 65 and over (male)
- Cervical cancer: 25-64 (female)

Although the real effectiveness of the methodology (which does make heavy data demands on local centers) must await results of the field trials, the examples cited in the preliminary report suggest it can be a sensitive and extensive reflector of health needs and of the interactions of a health system with them.

Outcome Evaluation and Child Health Care. There are several aspects of child health and health care services for children which make an assessment of the value of outcome evaluations different from what it might be for adult services. In general, possibilities for outcome evaluations which would be usable for policy purposes are in many ways

*Later we will discuss an HEW Program Analysis of maternal and child health care programs which is a starkly contrasting example of undocumented procedural issues and unsubstantiated value judgments, although its general intent and even conclusions are similar.

better than average. The major reasons for this optimism are discussed below.

1) Infant mortality (along with maternal mortality) is not only the best general community health indicator we have, but it is a direct indicator of risk to children--both of death and of impairments and deficits only marginally non-fatal (e.g., severe brain damage, extreme prematurity). (Knobloch & Pasamanick, 1966; Pasamanick & Knobloch, 1960, 1966; Birch & Gussow, 1970; Shapiro, Schlesinger & Nesbitt, 1968)

2) The measures of low birth weight and newborn Apgar scores (Apgar, 1953) can be unambiguously made and appear to have high correlations and causal linkages with health deficits of various kinds (Birch & Gussow, 1970; Drillien, 1964, 1961; Glass, Kolko & Evans, 1971; Illsley, 1966; Robinson & Robinson, 1965; Schachter & Apgar, 1959). However, a reliable direct measure of prematurity instead of the proxy measure low birth weight would improve the precision of the relationships found. For example, the very careful epidemiologic study done on all pregnant women and their offspring for 10 years on the island of Kauai (Werner, Bierman & French, 1971) shows that there are distinct differences in health status and health and behavior at 10 between low birth weight full-term babies, premature (i.e. not fully developed) babies of any weight, and babies of extreme low birth weight (less than 1500 grams). It is the last group which shows the most severe and irremediable handicaps (in health, size, and even IQ).

3) Morbidity in infants and young children is frequently due to infectious diseases which, because of their acute characteristics, are the most accurately measured type of morbidity, given current morbidity

indices. (AAP, 1971)

4) It also appears that dental health deficits can be quite accurately predicted by the simple negative measure of lack of an annual visit to a dentist. (Richmond & Weinberger, 1970) Furthermore, in carefully designed and controlled studies, fluoridation of public water supplies has been shown to consistently reduce the numbers of caries in children under 18 by 50 to 60%. (Ast, 1962; Arnold, et al., 1956)

The major problem in child health outcome evaluation is that other critical aspects of child health, functioning, and development, which are clinically obvious and publicly recognized, have not proven quantifiable with the same degree of reliability. In particular, chronic conditions, malnutrition, child abuse or neglect, and sensory/neurological deficits are difficult to measure and to relate to health services received.

There are still many unmet problems in the field of child abuse and neglect, for even the terms are not yet clearly defined. (Rowe, et al., 1970, p. 952)

The present nomenclature of the social illnesses of childhood describes groups of manifest symptoms; it defines neither pathophysiologic mechanisms nor the specific causes of disease because little is known about them...Clinical practice, to the extent that it is documented in the literature, seems to bear no consistent relation to case outcome for this group of illnesses. (Newberger, 1971, pp. 4, 8)

Except for data from studies of congenital malformations, little hard data exist on the incidence and prevalence of chronic conditions in childhood. Populations examined (and conditions counted) in the major studies are usually not comparable. Neither are definitive data available on the extent of disability caused by chronic illness. (Wholey, undated, p. II.6)

As noted in the definitions of outcome evaluation, only "measurable aspects of health" can be evaluated. While it probably is a good thing to go ahead and measure and assess all those elements of health which

one is able, there is a grave danger in so doing that those variables which are supported by quantified data fare better during program analysis and funding. While the approach of Wholey's group is a hopeful sign that clinical evidence and professional judgment are not slaves to evaluation data, the fact remains that there is much more money these days in maternal and infant care than there is in nutritional programs, chronic disease detection and treatment, prevention of child abuse, or the treatment of handicapping conditions and sensory and neurological deficits not prevented by prenatal care. The case remains intuitively open that in fact certain patterns of diagnosis and care do make a difference in the sequelae of these problems, but that we are not able to document them. It is also possible, of course, that we don't know what to do about child abuse or chronic handicaps and that no program would be successful, even if we could measure the outcomes. *

Impact

Impact evaluation is an attempt to extend the concept of outcome evaluation to encompass the psycho-social and socio-economic variables that influence health status and the success of health programs. This concept, advocated, for example, by de Geyndt, Weckworth, and the Systems Development Project, deals with the "impact of the health services system on the lives of the recipients and on their contribution to society. This means thinking in terms of total target populations or total communities." (de Geyndt, 1970, p. 33) From a different perspective, Suchman (1967) uses the term to mean the degree to which effective performance (effectiveness being a measure of efforts meeting stated objectives) is adequate to

the total amount of need. The concept corresponds to Freeman and Sherwood's (1965) term "accountability," and Stufflebeam *et al.*'s (1971) terms "importance" and "scope," all defined previously.

In fact both perspectives are closely interrelated. Any attempt to focus on the effects of programs on people must address the question of what difference does the program make. There are several ways in which programs can make a difference; impact evaluation is a method of analytically separating those different levels of effect. For example, a program to provide intensive care for high-risk newborns is medically innovative and successful in 80% of the cases it handles, but makes no dent in national infant mortality figures since it can cover so few children because of its cost and demands for the technologic capacity of a major medical school complex. On the other hand, another program may concentrate on providing only traditional, basic prenatal care, nutritional counselling and perhaps child care classes for pregnant women but emphasizes outreach and community participation in order to find the high-risk mothers who rarely obtain care until delivery. It is likely that much more medical research money and professional interest will be put into the first program without public intervention, since it is more suitable to the needs of the delivery system itself: staff positions for professionals, research opportunities, high and clear-cut "cure" rate, medical control. But if the impact of programs on social needs becomes a criterion, then the implicit trade-off between less spectacular (but positive) improvements for many people versus miracle cures for a few is brought out

explicitly.*

Impact evaluation cannot resolve these issues since they involve very basic value-judgments and philosophic beliefs on the mutability of human behavior, the role of expertise and scientific research in society, and so on. But it can provide fuller pictures of the programs to facilitate such judgments.

Impact evaluation need not operate only at this rather lofty, above-the-fray level. One well-developed model of an impact evaluation methodology is currently being used to analyze the gaps in the system of child health care, i. e. to measure the real versus the ideal according to some fairly well-agreed-upon criteria. (Minnesota Systems Research, Inc., 1972) Those criteria, in the alliterative style popular among health planners, consist of six A's:

- 1) Appropriateness. This criterion has two dimensions: a measure of need (measured by proxy indicators of risk factors; see infrastructure discussion, supra.) and program elements with priority rankings possible (e.g., preventive, diagnostic, therapeutic, rehabilitative, educational)
- 2) Availability. This criterion contains factors of
 - eligibility
 - services, by eligibility category
 - standards
 - financial mechanisms

*It can also be seen as a trade-off between curative and preventive medicine or as between medical care and health education in a broad sense. Another example which highlights these multi-dimensional policy issues is the contrast between investment in expensive (but life-saving) heart-lung machines and heart transplant teams versus massive research and educational campaigns on the ties between heart disease and smoking, air pollution, eating habits, and other social variables.

- 3) Accessibility. This criterion examines
- how care is delivered
 - sequencing of service (to what degree is continuity of care reached within the program)
- 4) Acceptability.
- to seekers of care
 - to providers of care
 - to payors
- 5) Accountability. This criterion evaluates two forms of accountability:
- dollar expenditures
 - accomplishment or outcome of program (such accountability being shared by creators, implementors, and users)
- 6) Adaptability. This criterion is the dynamic linkage of the first five criteria over time.

The first four A's are seen as a dependent sequence, "a sequence of progressively more involved commitment to specifying programs of service." (MSR, Inc., 1972, p. 47) This concept of sequencing evaluation, i.e. looking at appropriateness before studying availability, etc., is a major innovation. For example, it recognizes the potential value of utilization data, but only within a previously specified context:

Within the constraints of both availability and acceptability, the utilization of the services of programs, when choices do exist, is the ultimate reflection of acceptance. (MSR, Inc., p. 46)

In many ways, impact evaluation incorporates many of the components of evaluation already mentioned--content, process, organization, and outcome--into a conceptual framework relating them to each other. A determination of appropriateness assures some correlation between health needs and program elements and hence requires prior outcome evaluation studies or assumptions. Acceptability factors are clearly process variables. Accountability, while very sketchily developed by MSR in the design,

could include both cost-effectiveness studies and efforts to allocate professional and patient responsibilities for health care and health, thus enlarging the parameters of traditional auditing, PPBS, or cost-benefit concerns.

The major independent significance of impact evaluation is the measuring of program effect in relation to total need. This presumably helps avoid elegant, drop-in-the-bucket program designs; it is also a critical input for any sort of comprehensive health planning, where the major concern is in summing up the totality of health and health-related activities in an area and comparing them with some picture of the total health needs of the community. (c.f. Kerr and Trantow (1970) for a fuller explanation of this use of assessment or evaluation) In this sense, many reviews of "present strengths and weaknesses in current systems of comprehensive health services for children and youth" (Haggerty, 1970a), and such are really informal impact evaluations of the health system as it serves a particular group. Haggerty's definition of adequate care, which he equates with comprehensive health care, includes many of the same criteria (and in much the same order) as MSR's: "complete, competent, continuous, coordinated, compassionate, and community-oriented." (p. 74) He implicitly links the importance of impact evaluation to the increasing stress on preventive care and mentions the American Academy of Pediatrics' goal of concern for the welfare of all children. These make it "necessary to identify the population for which a physician is responsible, especially those at high risk, and ensure that they receive care." (Ibid., p. 84) He quotes Dr. Kerr L. White's comments at a conference on child health evaluation:

This shift in concern and responsibility from that of the individual clinician confronting the individual patient to the concern of all clinicians for whole communities or large groups of physicians for large groups of individuals represents a fundamental shift in orientation in medical education and practice. In epidemiologic terms the denominator changes from that of the physician's practice, or the hospitals' lists of visits and admissions to that of a population in a catchment area, a geographic territory, a prepaid group insurance scheme, or a community. (In Haggerty, 1970a, pp. 84-85)

Impact evaluation, then, is basically a means of making evaluation studies directly relevant for and more transferable to planning and policy-making.

Advocacy

Advocacy is used here to mean those goals and objectives to which health programs are addressed, both explicitly and implicitly. This is an area which is rarely analyzed in discussions of program evaluation. When it is mentioned, the aspect dealt with is the ambiguous or contradictory nature of stated program intent, in legislation, in guidelines, or in practice. Since evaluation of any kind requires stated objectives of some sort as a focus, there are immense difficulties in designing an evaluation study for federal programs whose enabling legislation is usually vague and the result of compromises of various types to begin with and whose nature is then altered, often beyond recognition, through the process of administration and regulation. Thus it should be part of any program evaluation to trace the legislative history of the program, to identify some of the hidden agenda items to which it was addressed, to relate those goals or objectives to actual priorities as reflected in agency guidelines, and to relate those priorities to the needs, objectives, and constituencies which the projects in the program

themselves feel that they are addressing. Another function of advocacy studies should be to examine the extent to which the public goals of the program were not meant to be fulfilled. (c.f. Etzioni, 1962) In fact, a systems model of evaluation is in general much more likely to consider such questions as the real intent of goals, their interrelated context, and their functional reality.

Such an analysis is not only fascinating and instructive in its own right as a study of the process of social policy implementation, but it also is essential to a meaningful (and equitable) evaluation of program outcomes or process. One of the most valuable results of evaluation of advocacy is the making explicit of the non-health goals of purportedly health-directed programs, since, "the true purpose of a health program may be only peripherally related to health." (Thorner, 1971, p. 531) Elinson and Herr (1970), for example, characterize the neighborhood health center movement as largely a political and social reform movement, not a health movement. The latent objectives they found included:

- 1) improving the image of the black male in poverty communities;
- 2) stimulating and maintaining solidarity among migrant Chicano farmworkers;
- 3) pacification of hostile communities by colonial powers;
- 4) discharging missionary service obligations of the medical-hospital establishment;
- 5) filling a political void in social and economic action; and
- 6) politicization or radicalization of youth. (pp. 98-9)

Another hidden agenda frequently found is an economy motive to save public funds. (c.f. Thorner, 1971) Much of the current fuss about

national health insurance, HMO's, and peer review is due to a growing crisis in medical care prices and federal and private expenditures for health. Such economic objectives need not be at odds with health ones. Prepaid group practice appears to have many values in terms of quality of care and patient satisfaction as well as its demonstrated effects on hospitalization rates and premium prices. (Donabedian, 1969) Early screening and treatment for handicapping and chronic conditions in children not only may save the state substantial investment in later hospitalization, institutionalization, and special education, but also enables those children to either escape serious handicapping or to grow up as more completely independent and normal adults.

On the other hand, national nutrition programs, which could be advocated for the alleviation of hunger for children and families, are administered by the Department of Agriculture and have been generally advocated instead as farm-support programs. The concern of the Department with not allowing the programs to in any way enable people to get free what they would otherwise have purchased has meant that the programs, particularly food stamps and commodity distribution, have always been more concerned with fluctuation in agricultural prices and keeping costs down than with feeding people. (See, for example, Segal, 1970; Kotz, 1969; Citizens' Board of Inquiry, 1968, 1972; also, discussion of nutrition programs in Chapter 4.)

Another dimension of advocacy, one even more neglected in the literature, is the correspondence between program goals and objectives (diverse a set as they may be) and other national, local, and individual goals and priorities. Fanshel and Bush (1970, p. 1022) seem to have this

in mind when they formulate complete performance analysis as including a measure of conformation to "society's rules of ethical and rational behavior."

This measure is clearly related to impact evaluation: one aspect of correspondence is the extent to which government programs, singly or together, are an adequate and appropriate response to total social needs. This is what programs are "accountable" to.

This approach also enables one to assess whether, even though individually effective, a program may not be counterproductive in the light of other concurrent efforts. For example, from the standpoint of a community with tremendous perceived health needs and desires to meet them, the current array of categorical federal programs with incompatible eligibility, reporting systems, and control requirements seems not so much beneficent as an example of irrational program planning. Similarly, as the social values of "health care as a right," consumer participation or community control, and comprehensive care become more widely accepted, existing and planned programs need to be examined in light of those values, even if not designed to directly implement them. In this sense, advocacy questions can serve to point out anachronistic or inappropriate policy philosophies as the social and political situation changes.

And finally, Fanshel and Bush's measure provides a framework in which one can directly face the potential conflicts between society's ethical norms and its rational ones. Why should food programs be designed to serve the poor rather than the farm lobby? Or should they serve as health care programs, under medical direction? How do we deal with program goals which express moral commitment (at least on paper) but have

no clear, rational way of being administratively carried out?

The best examples of evaluations of advocacy which I have found have been by people in the journalistic and radical muckraking school, not by health planners, policy analysts, or program directors. For example, there are several excellent analyses of hunger in America which clearly reveal the hidden agendas of the federal agencies and the failure to define the problem in ethical terms. (Kotz, 1969; Segal, 1970; Citizens' Board of Inquiry, 1968, 1972) In a more provocative and less well-supported argument, The American Health Empire: Power, Profits, and Politics (Ehrenreich & Ehrenreich, 1970) examines the non-humanitarian side of "big health": insurance companies, medical schools, and government programs.

Actionability

Actionability refers to the vitality of a program as a program. It is a measure of viability, of flexibility, of life force. It is an ambiguous measure, one which I have formulated because this dimension of program effectiveness is so rarely addressed directly or seriously in evaluation; or else it is presented as the entire story in "why-we-went-wrong" articles. This is a true systems-oriented measure, which ideally could combine those two extremes.

Actionability is in some ways very similar to advocacy evaluation: a program clearly at odds with prevalent cultural or political trends (e.g., federally funded abortion clinics or serious massive attempts at community control of health facilities) does not stand much chance of survival or success, regardless of the logical and medical soundness of

its structure and operation.*

Actionability, however, includes several other factors:

- funding: its adequacy and stability; strings attached
- factors outside program control: their power, areas of impact, chances for by-passing or winning over (e.g., local board of health vs. free clinic; medical society-controlled licensing laws vs. new professionals training programs)
- political climate (e.g., election year hesitancy to innovate or raise budgets; dismantling of OEO under hostile Republican administration)
- internal feasibility of achieving objectives (e.g., expecting medical schools to work closely with community groups in co-sponsoring a program without extensive preliminary preparations)

In other words, it is not fair to judge a program a success or failure without examining the environment into which it was placed and the basic nurturance it was given. This component of evaluation can be seen as an attempt to evaluate programs as one input into an existent system. Weiss and Rein (1969) have characterized this approach as a "non-experimental methodology for evaluation research" appropriate for broad-aim social programs. They call first for a more descriptive, inductive methodology, concerned more "with learning than measuring;" the approach has much in common with the case study method of evaluation and analysis. Second, they note the systems perspective which can address "such issues as the way in which the program makes a place for itself, the new stresses it introduces, and the way the system accommodates itself to the program, as well as...the issue of what individual and institutional benefits the program brought into being. (Weiss and Rein, 1969, reprinted in Caro,

*The recent Supreme Court ruling on the unconstitutionality of anti-abortion laws, however, may have radically shifted the potential actionability of federally-funded abortion services.

1971, p. 296)

Not only can such a systems perspective provide a richer and more dynamic understanding of innovative health programs, but it can also increase their influence and ability to survive political or cultural battles, in which they rarely figure as principle figures.

It is obvious that the decisions are made in the political arena, in a broad social and economic context, influenced by unstated objectives often unrelated directly to health. These program objectives, if recognized and overtly stated, could be evaluated by appropriate techniques. However, an evaluation which is constrained solely to the effects of the program on health can be expected to have only a partial influence on the decision process, and quite often a minimal influence, depending on the relative strength of the influence of the nonhealth factors. (Thorner, 1971, pp. 531-2)

Of particular importance is a consideration of how a program fits into and changes the administrative structure to which it is attached. As Williams and Evans have commented in an analysis of evaluation of Head Start:

In thinking about the development of evaluations, it must be remembered that after a decision is reached, the further hurdle remains of translating the decision into effective operating policy so as to improve the performance of the agency's programs. Those who plan evaluations need to be sensitive to an agency's administrative structure through which policy decisions are implemented for, in the final analysis, the test of the effectiveness of outcome data is its impact on implemented policy. (Williams and Evans, 1969)

An excellent example of the difficulties of survival of even the most carefully researched, documented, and agency-supported health programs is that of the major program analysis on child health carried out by a high level HEW task force in 1966 (HEW, 1966; Wholey, undated; Levin, 1968). In reviewing the fate of this analysis, designed to be both directly usable for and convincing as policy, Wholey notes that while the 1967 Child Care Act proposed by President Johnson and later rewritten by Congress still included the major elements recommended

by the program analysis group, "the new legislation was not supported by increased fiscal year 1968 funding....In a very tight budget year, Congress took a general position of not funding new programs." (Wholey, undated, p. III.4) Even more discouragingly, "by fiscal year 1970, except for continued emphasis on family planning programs, the new thrusts in the Child Health Act of 1967 had largely disappeared from the Administration's program requests." (p. III.5)

Finally, we should briefly mention the importance of actionability (and advocacy) evaluation as a way to analyze and evaluate evaluation studies themselves. It should be clear from the vast array of issues mentioned in this discussion of evaluation, and with the great difficulties and limitations associated with each, that no one program could possibly afford to do this type of complete evaluation. Decisions need to be made on when, where, how, and by whom evaluations should be done.

[One] way to look at evaluation is as a programming input, which may be subject to evaluation just like other inputs. In cost-benefit terms, the cost of evaluation should be related to the benefits that evaluative data and judgments contribute to programming efficiency or effectiveness. A heavy investment in formal evaluation is most likely to be justified when a program is expensive, when its impact is potentially great but uncertain, and when there is a great potential for diffusion of programming concepts. (Caro, 1971, p. 7)

Rossi (Chapter 2 in Rossi & Williams, 1972) advocates program evaluation, especially with a research design, when the rates of change in outcomes are expected to be rather small and thus careful study will be needed to demonstrate them. Thorner likewise suggests that studies of outcome or effectiveness should be carried out in carefully selected situations, with concern for transferability of results to other places and situations, "so that end results in service programs can be inferred from the study

of process or intermediate goals, in the manner indicated by Shapiro (1967)." (Thorner, 1971, p. 532)

The HEW program analysis group found in 1966 that the "Federal Government had no system for managing evaluation of the effectiveness of its programs...there was no system for planning what research and evaluation studies were required to measure the overall effectiveness of Federal maternal and child health programs or the relative effectiveness of different projects within these programs; and there was no system for review and use of program evaluation and project evaluation studies in reaching policy decisions." (Wholey, undated, p. IV.1)

Despite the subsequent beefing up of the Office of the Assistant Secretary for Planning and Evaluation and the increase in evaluation funds directly under its control (via the earmarking of up to 1% of appropriated funds under several pieces of legislation), there is little evidence that the situation is substantially different today. (For an excellent and more optimistic discussion of the growing HEW coordination of planning and budgeting see Iglehart, 1971.) One important attempt to bring together a comprehensive review of federal evaluation efforts and recommendations toward development of an actual "federal evaluation policy" is the recent work by Wholey and others (Wholey et al., 1972). Their model has been discussed earlier. The ultimate usefulness of these recommendations, however, is limited by their overly optimistic, rather narrow, PPBS-oriented attitude toward evaluation methodology. The review of the politics of evaluation and general recommendations for more planning, funding, and accountability for evaluation are very pertinent, however.

This concludes the discussion of my tentative framework for child health program evaluation. The components of advocacy and actionability are still only sketchy and need further abstraction and clarification. Some of the other more traditional components need reduction and reassessment. And especially, the whole framework needs to be applied, either to sets of previous evaluations to see if their results can be synthesized or to new and better designed evaluations to see how far this technique can actually go in deepening our understanding of social processes and in shaping our policy decisions.

In the next chapter, I return to the world of child health to examine the needs and knowledge which will shape those future evaluations.

Chapter 4.

Needs and Knowledge: Critical Child Health Problems and Existing Approaches to Solve Them

Part 1.

Health Needs of Children

It may seem that I have strayed a long way from "planning child health programs." Indeed, the preceding extensive examination of the nature and limitations of evaluation may make planning seem like either a hopeless or an arrogant activity. Evaluation, however, is only a synthetic activity. Behind it there is the reality of world of children, both healthy and sick, and of a multitude of activities, both deliberate and inadvertent, which affect these children. Ahead of it is the reality of policy which does get made despite the lack of data, laws which get passed for anything but scientific reasons, and the actions and lives of us all which seem to go on whether they are evaluated or not. In this chapter, the reality of child health needs and programs will be introduced.

As viewed here, the prenatal period and the early years of a child's life represent a series of "critical periods" in physical and psychological development during which he or she is susceptible to risks of impairment which may confer lasting disadvantage. The task of minimizing those risks belongs partly to the family, partly to professional health services (private or public), and partly to non-health services (income, housing, safe air, safe highways, etc.).

My concern will be primarily with organized health services. It should be noted here that "critical periods" can reflect not only risks with permanent consequences but also those whose long term effect is negligible or unknown, but which are undesirable in their own right and tend to occur at certain periods in the child's life. For example, broken arms or legs rarely cause residual or permanent damage but that is no reason why public policy or families should not be concerned about the safety of playground equipment or with the adequacy of emergency orthopedic services. Only the rationale for the concern will be different.

Child Development and Child Health

This distinction between consequences is important because of current emphasis on child development as an organizing theory to conceptualize child needs and programs. In the Introduction, I raised the question of whether child development was a useful or limiting concept in evaluating and planning child health programs. Without going into any detailed examination of child development, I would like to suggest the outline of the argument here.*

The issues of physical and psychological development are quite distinct in some senses, completely intermingled in others. The distinction is fostered by the fact that different knowledge, beliefs,

*This argument is drawn from research I did for the Huron Institute (1972), whose report, done under contract to HEW, is a thorough compilation of our current knowledge about child development and programs successful in nurturing it. The importance of the concept is shown by the fact that the Government is currently funding another major study focusing on child development by the National Academy of Sciences.

professions, and technologies are involved in dealing with them. Different aspects of parental behavior are also involved in the caretaking functions in each area. Despite these differences, there is a strong present trend to mesh the two as child development, and to design programs that can respond comprehensively to problems of either physical or psychological development. There are some good reasons for these efforts, because quite often both kinds of support are needed in order for either intervention to be effective. Psychological neglect of the child will often eventuate in health problems: problems of malnutrition, plumbism, failure to thrive, abuse, or emotional disturbance. Physical problems -- various kinds of handicap, mental retardation, chronic illness -- may make heavy demands on the family's psychic care of the child and on the child's social development. Because of this interrelation, there has been a tendency in the last decade to try to construct programs marked by "comprehensiveness" or "coordination", to bring together services in education, health, and family work, and to stress "child development" as an integrating concept.

Richmond & Weinberger (1970) present a concise history of the emergence of child development as a focus in pediatric medicine. Before 1900, pediatrics did not exist as a medical specialty. From 1900 to about 1925, pediatrics was in a "descriptive era of pediatric nosology." An era of laboratory investigation, bringing with it emphases on specific etiology and therapy and the first child development institutes began in 1925, to be followed in the early 40's by an era of pediatric therapy. Finally, about 1960, an era of child development (based on developmental biology) and of prevention began

and continues today.

In the development of health services, a system of priorities emerges. In most cultures, the primary concern is with physical health; mental health programs generally must await the emergence of sound general health and welfare services. As the physical health of children has improved, there has concomitantly developed an expectation that child health workers will concern themselves increasingly with the cultural, social, and psychological aspects of child care. (Richmond & Weinberger, 1970, p. 24)

Because child development attempts to integrate parts of education and medicine, it is important to point out some differences between these two disciplines and the contexts in which they operate which have policy implications.

The primary goals of virtually all educational programs are developmental. More accurately, the purposes explicitly stated and evaluated are developmental in nature. Health programs, either in general or for children, are concerned primarily with alleviating or preventing conditions which, whether or not they have permanent developmental consequences in a cognitive sense, are deemed undesirable--either for the individual or the society--and are considered either curable or manageable. As a result, discussions and studies of the developmental implications of various health problems (malnutrition, lack of prenatal care, etc.) tend to be peripheral rather than central to the core of health care literature and research. What literature does exist tends to be less voluminous, more testimonial in nature, less replicated, and more idiosyncratic than the corresponding literature in educational and behavioral science fields.

This trend is reinforced by the greater scientific "hardness" of other health issues, such as the minimal dosage of medication to

achieve results, development of simple but reliable tests for diseases and conditions, precise surgical techniques, or monitoring and life-support technologies. Whether this is bad money driving out the good or not depends on one's values, but the fact remains that developmental issues (except in strict anthropometric or physiological terms) and psycho-social issues in general are not widely accepted as completely proper topics for serious medical research. (For an excellent development of the counter-argument for the necessity of behavioral science, an argument gaining strength in medicine, focused directly on children, see Talbot, Kagen, and Eisenberg, 1971, passim, and particularly Talbot and Howell, 1971.)

Most educational situations are focused primarily on children or child-centered issues. That is, experts in the educational field all accept a concern with the child and his development as most proper, even if they disagree about goals or methods. In health, however, pediatrics, child psychiatry, and even public health nursing (a major source of preventive and maternal programs) are specialties within much larger professional domains. It is not, therefore, possible to take for granted an uncomplicated dedication to child health in those who will provide it, as it is just one part of their responsibilities.

Finally, while education at the elementary level (and increasingly at the preschool and day care levels) is overwhelmingly a public function, health care is not. Public health efforts are in general kept quite separate from other health activities and are definitely something of an orphan in the medical professional and status hierarchy. Even today, seventy per cent of all expenditures for child health care

services are paid for by private means (out-of-pocket expenses or through insurance plans), while twenty per cent are paid by state and local governments, and only ten per cent by the federal government. (Breslow, 1969, p. 328) There are several important implications of this private, marketplace nature of health care delivery:

- a) Governmental leverage is likely to be less than or at least different from that in, say, public education.
- b) The context of public programs and the independent variables which they can affect are often quite different from those in educational programs.
- c) The health care "system" is less of a system than is the educational one. Except in some categorical programs, there are no state-wide mandates, uniform standards for quality of care or process organization, or reporting requirements. The potential for large scale evaluation studies with even minimally acceptable control groups is thus limited.
- d) Finally, the public accountability and visibility of education, which it could be argued is a primary reason for outcome evaluation efforts, is virtually entirely lacking in health care, particularly in so-called "mainstream" care delivery for non-indigent persons.

At the present time, there have been no substantive, major studies of the trade-offs between health programs and others specifically intended for child development. But we do know ways in which health or health care can interact with development and educational achievement, and some areas in which it seems probable that health and child development programs could represent trade-offs. Scientific evidence and logical sense suggest that the processes of physical growth and/or impairment are distinct and separable from processes of cognitive, emotional, and social development only in the formal schemes of

professionals in those fields. Some brief, acute childhood diseases or traumas (chickenpox, minor colds, broken arms) may be generally considered as isolated medical events with no lasting implications. However, for many children contact with the medical care system, with the sick role, or especially with even brief hospitalization in early childhood can have significant and often permanent effects on behavior, attitudes, and later illness patterns and response. (Bergman, 1965; Dimock, 1960; Spitz, 1945, 1946; Shaeffer & Callender, 1959; additional bibliography in Shore, 1967)

Other evidence, comprehensively reviewed and synthesized by Birch and Gussow (1970), strongly suggests that physical health is a critical independent variable underlying educational performance and intellectual development. (See also Richmond & Weinberger, 1970; Gussow, 1970; Callahan, 1971; Chase & Martin, 1970; Elias, 1971; Kaplan, 1972; Kerr, G., 1972; Read, 1972, 1969; Werner, et al., 1971) Assessment of biological risk, illness history, and current health status in groups of children under study for evaluation of schooling or cognitive stimulation programs would seem essential if the results are to be meaningful. And yet this is virtually never done, and the implications for rather gross false negative findings in education research appear startling. For example, Christopher Jencks' new and impressive reanalysis of the data on education and inequality still does not examine or control for the health or nutritional status of the children. This is not a trivial omission.

Children who had been repeatedly and excessively exposed to biological risk, both before and after birth, were unlikely to be dramatically helped solely by the application of "more schooling," no matter how early it was begun or how intensively it was pursued. (Birch & Gussow, 1970, p. 264)

Birch and Gussow go on to point out that the physiological and developmental impact of sub-optimal health or conditions of life cannot be wiped out in a short period of time, and thus evaluations may be further upset if they seek to draw conclusions about intervention programs after a year or even a few years. For example, the substantial evidence that the health and physical condition of the mother, particularly during her own childhood and adolescence, is one of the greatest risk factors for the infant indicates that one must wait until the next generation at least (and probably longer) to see the ultimate developmental or learning effects of improved maternal health status.

The evidence, however fragmentary, leads one to the conclusion that biological risks, unfavorable environments and poor health are not random occurrences in our society but rather tend to pervade certain groups in such a way as to create widespread and rather massive differences in development. Finding and analyzing such a framework would seem to be vital in eventually settling, among other things, the current debate over genetic potential as a causal factor in achievement differentials.

Thus, although it is entirely reasonable to wish to know more about the heritability of intelligence and learning capability in human beings, the pre-condition for a serious consideration of this question is the equalization of the developmental environments of those who genetic make-up we wish to compare. (Birch and Gussow, 1970, p. 266)

For example, comparisons in which current family income or social class are used as control variables, thus apparently equalizing environment, are simply not valid unless the past health and nutritional history of those groups were also roughly comparable. And even a brief look

at national income figures will show that very few black or other non-white children live in families who have been out of poverty for as long as their white middle class peers.

The area in which there is the strongest evidence of a trade-off effect is that of early screening and treatment for handicapping and chronic conditions. The Joint Commission on Mental Health of Children (1970) estimated that comprehensive care until the age of three could reduce chronic handicapping conditions by 20 to 30 per cent. These conditions, if not remedied, generally require special educational and/or physical rehabilitation services which are usually quite expensive and therefore difficult to provide for all children in need. An HEW Maternal and Child Health Program Analysis (1966) reached similar conclusions, placing stress not only on detection and treatment but also on preventive measures such as prenatal care and family planning. There is also evidence, summarized by the Huron Institute report (1972, Chap. 12), that early treatment in a setting combining health and education will reduce the level of handicap or lower achievement patterns of handicapped or retarded children. Not only do educational evaluators weaken their programs and findings by not measuring health status, but children who are "claimed" by the medical profession may not have proper attention paid to their educational needs. As Eliot Richardson, ex-Secretary of HEW, has said:

The dimension of the handicapped pediatric population is such that we can no longer isolate educationally what are also complex medical problems. Follow-up studies of children with cerebral palsy, myelomeningocele, epilepsy, and profound deafness illustrated that these children are seldom if ever prepared adequately for adult life, and they may become victims of society in the broadest

sense...left without guidance, and no gainful occupation...
to give the patient self-respect and a place in society.
(quoted in Haggerty, 1970a, p. 86)

The whole notion of preventive health measures, which are becoming very popular, is also that of a trade-off. Such programs as rubella inoculations for children, genetic counseling and testing of parents and fetus, family planning, and abortion services may--at reasonably low cost--reduce the number of unwanted, neglected, abused or handicapped children who would otherwise require special medical, educational, and probably family support services.

Normal Patterns of Child Health

Issues of child health need to be put in the perspective of national norms of health. Basically, childhood, after the first year, is normally the healthiest portion of an individual's life. The level of those norms of health are dependent upon the society's general level of development, its investment in health resources (medical schools, hospitals, biomedical research), and, as evidence will show, the degree of income and standard of living disparities in the society.

The American Academy of Pediatrics, in its excellent review of child health needs and programs, Lengthening Shadows (1971), summarizes the acute health problems of children (those lasting no longer than three months) by separating them into four chronological subgroups:

- 1) Neonatal Period* (0 - 1 mo.). During this period the child is at great risk from many acute conditions, some the result of perinatal complications or congenital defects, and some due to the environment--the newborn nursery or the home. Among the neonatal conditions are infections acquired prenatally (rubella, congenital pneumonia), during birth (herpes virus), and postnatally (staphylococcal infection, pyoderma, sepsis, pneumonia, meningitis, epidemic diarrhea of newborn). Other conditions requiring quick diagnosis and treatment are hyperbilirubinemia (which can either kill or cause brain damage), biochemical disturbances (such as hypoglycemia and hypocalcemia), severe anemia, respiratory distress syndrome (still a frequent killer), and malformations which may require surgical intervention.
- 2) Infancy (1 mo. - 2 yrs.; this does not correspond to the strict use of the adjective "infant" for mortality rates which refer only to first year). This is the period of most acute illnesses while the child is also achieving greater control of immunological, physiological, and biochemical functions. Some of the common acute conditions may cause permanent impairment (or death) if not prevented or treated speedily: acute respiratory infections, otitis media (frequent cause of permanent hearing loss in medically underserved areas), acute infantile diarrhea, pneumonia, pyoderma (if not treated), and acute bacterial meningitis.
- 3) Preschool (2 - 6). Acute illnesses occur less frequently in this period, partly because of the child's immunities developed earlier. Acute contagious diseases (as the child tends to mingle more with his peers), accidents, and poisonings are all at peak frequencies, however.
- 4) School period and adolescence. Children are generally in better health than earlier. Chronic diseases (rheumatic fever, diabetes, etc.) often begin to appear now, but at incidence rates much lower than childhood acute conditions. Accidents are by far the leading cause of death, followed by cancer.

*Some definitions of the subdivisions of infancy may be useful:

prenatal: conception - birth

neonatal: birth - 1 mo. (28 days)

early neonatal: birth - 6 days

late neonatal: 7 days - 28 days (or 1 mo.)

postneonatal: 1 mo. - 11 mos.

infancy: under 1 year

perinatal: period surrounding birth; depending on study, may begin with 20th to 28th week of pregnancy and extend from 1st week of life (early neonatal period) to 1st mo. (neonatal period).
(Sartwell, 1965, p. 657)

By definition, chronic conditions and handicaps do not display the same patterns of age-specific immunity or frequency that acute conditions do. Patterns seem, however, to be pointing toward an increase in the number of children whose condition is congenital or from birth, rather than acquired during childhood (e.g., congenital blindness vs. blindness by trauma) and whose problems are multiple. There is some indication that prevalence patterns in later childhood may be altered by screening and treatment in the early years. (Joint Commission on the Mental Health of Children, 1970; Maternal and Child Health Programs, 1966)

In 1965, approximately one in every twelve children (aged 6 - 11) had a speech defect; one in nine had defective vision; and less than one per cent had hearing problems. The incidence of blindness among children under 20 has remained steady since 1960. (White House Conference on Children, 1970, pp. 30-31; see also AAP, 1971, pp. 51-27; Richmond & Weinberger, 1970, pp. 37-39) While techniques for preventing and detecting handicaps at an early age (such as amniocentesis, preventive inoculations, Rh vaccine) are rapidly becoming universal, there are other factors increasing their likelihood: expertise in hospital deliveries which saves many infants who would previously have died and the ever growing chance of iatrogenic or genetic effects from the drugs administered to or taken by women during pregnancy. (The average number taken is already four; Richmond & Weinberger, 1970, p. 39.)

Critical Factors for Child Health

In this section, I shall be looking at need from all of the perspectives outlined in Chapter 1: as individually discovered "crises", as conditions amenable to treatment by best available expertise, and as systematically analyzed gaps or misfits in service programs. There is a surfeit of information now available on the subject, from all three perspectives. Most of my own research and reading has been in the third category, with forays into the others.

The organization of the material has a three fold purpose. First, to provide some substance to abstract models of evaluation: to suggest those areas most in need of attention in order that the feasibility of obtaining adequate evaluations in those areas (according to the constraints laid out in Chapters 2 and 3) could be assessed. Second, to demonstrate in its arrangement and presentation my own attempt to force scientific evidence and moral commitment as close together as possible. And finally, to persuade the reader that the planning and execution of child health programs, even in the face of incomplete knowledge, is neither a trivial nor an escapable task for social policy in the seventies.

In the pages that follow, I first present in tabular form (Tables 1 and 2) what I have found to be either 1) the most commonly discussed or 2) the most critical in impact of health-related risks to children (primarily children under 15). I make a preliminary distinction between independent risks (Table 1) and clustered ones (Table 2). Independent risks happen either individually or in random combinations to children (and to their pregnant mothers). Clustered risks, although not all

Table 1.
Independent Risks*

A. Biological Risks

1. Genetic

hemophilia
 sickle cell anemia
 male (across all income and racial groups, males have
 higher infant mortality rates than female babies!)

2. Congenital errors (Rutter, Tizard, & Whitmore, 1970)

malformations
 mental retardation (of certain types)
 cerebral palsy, other physically handicapping diseases

3. Maternal condition

nutrition (Scrimshaw, 1950)
 iatrogenic effects (from tranquilizers, fertility pills,
 etc.)
 chromosomal damage (from smoking, LSD, etc.)
 congenital syphilis or heroin addiction

4. Reproductive complications (Schachter & Apgar, 1959)

5. Diseases (C.f. previous AAP categories for age specific
 breakdown)

malignant diseases
 childhood diseases for which no vaccine is available
 acute respiratory syndrome
 sudden infant death syndrome
 rare diseases
 diseases of the teeth

6. Prematurity (of some types; Drillien, 1964)

* General references which have discussions of many of these risks are Lengthening Shadows (AAP, 1971, pp. 10-88; 152-191); Profiles of Children (White House Conference on Children, 1970); "Critical Health Needs of Children" in Chapter 6 of the Huron Institute report (1972) (which incidently I did not write!); Richmond and Weinberger, 1970; Haggerty (1970 a,b); and HEW (1966).

B. Social/Environmental Risks

1. Accidents (the leading killer up to age 15; 1/3 of all deaths) (Wight, 1969)

 motor vehicles
 drownings
 fires/explosions
2. Nutrition (Kerr, 1972; Scrimshaw, 1965; USDA, 1968)

 Monosodium glutamate, other dangerous additives and preservatives
 empty calories ("the Pepsi generation")
 obesity
3. Child abuse/neglect (Gil, 1970; HEW, 1968; Kempe, 1962; Newberger, 1971; Silver, et al., 1959; Simons, et al., 1966)

 failure to thrive (Talbot & Howell, 1971; Talbot, 1963)
4. Mental illness (NIMH estimates that only 5% of children needing psychiatric care are receiving it; WHCC, 1970, p. 29) (Joint Commission on the Mental Health of Children, 1970; Rutter, Tizard, & Whitmore, 1970)
5. Abortion: to the extent that abortion is based on knowledge of risk and reduces number of children born with congenital malformations, etc.

Table 2.
Clustered Risks*

A. Biologic Risks

1. Maternal condition (Yerushalmy, 1938; Yerushalmy, et al., 1940)
 - general poor health
 - poor nutrition (anemic, etc.) (Scrimshaw, 1965; Tompkins, et al., 1955; Committee on Maternal Nutrition..., 1970)
 - high parity
 - age (teenage; over 30) (Menken, 1972)
 - spacing of children
2. Reproductive complications (Knobloch & Pasamanick, 1966; Pasamanick & Knobloch, 1960, 1966)
 - toxemia
 - anemia
 - miscarriages
 - prematurity (Drillien, 1964; Tompkins, et al., 1955)
 - postnatal asphyxia
3. Congenital errors
 - malformations
 - mental retardation (Kaplan, 1972; Read, 1969, 1972)
 - sensory deficits (Vernon, 1967)
4. Extreme low birth weight (under 1500 grams) (Drillien, 1961; Glass, et al., 1971; Robinson & Robinson, 1965)
5. Infectious diseases (if generally controllable: pneumonia, influenza, diarrhea, gastroenteritis)
6. Cavities; missing teeth
7. Infant mortality (Hunt, D., 1969; Hunt, E., 1970; Shapiro, et al., 1968)
8. Maternal mortality
9. Nutrition (Chase & Martin, 1970; Christakis, et al., 1968;

*The general references most useful here were Birch & Gussow (1970); Gussow (1970); Richmond and Weinberger (1970); American Academy of Pediatrics (1971); Hellmuth (1967); HEW (1966); Werner (1967); and Werner, Bierman, & French (1971).

Citizens' Board of Inquiry, 1968, 1972; Fomon, 1970;
Gutelius, 1969; Kaplan, 1972; Owen & Kram, 1969;
Sandstead, 1971; Scrimshaw, 1965; Woolsey, 1971)

iron-deficiency anemia
malnutrition
hunger

B. Social/Environmental Risks

1. Non-medical status (Wagner, 1970; North, 1967)

poverty (Rider, et al., 1955; Sandstead, 1971)
non-white (esp. black, Indian, Chicano)
(Adair & Deuschle, 1970; McDermott, et al., 1972)
rural or inner city
migrant (Siegal, 1966)

2. Accidents (Wight, 1969)

lead paint poisoning, ingestions (Smith, 1964)
traffic
falls

3. Infections

unsafe water
inadequate heat
rat bites
poor nutrition (Scrimshaw, 1965)

4. Medical system gaps (Collver, et al., 1967; de Vise, 1969; Lesser, 1965; Madison, 1969; Monahan & Spencer, 1962; Strauss, 1967; Yerby, 1966)

no prenatal care
no well-baby care
no regular dental care
missing immunizations
unattended illness
no health insurance

5. Psychological/behavioral correlates (Callahan, 1971; Elias, 1971; Gussow, 1970; Read, 1969, 1972; Kaplan, 1972; Sandstead, 1971)

poor school performance
inattention
apathy
lowered intelligence, cognitive ability

proven to be dependent or linked causally can be shown to occur non-randomly. Independent risks are not synonymous with chance occurrences. A mother who is a carrier for hemophilia will have hemophilic sons; one who is not a carrier will not. But there is no evidence that that mother will also bear retarded or premature babies or that she will be unable to feed them adequately. Similarly, families who abuse their children may present clear psychiatric syndromes (abuse by their parents, extreme stress, uncontrolled anger, etc.), but those syndromes do not predict rates of tuberculosis or cerebral palsy in the children. For this reason, the entries on the "independent risks" charts bear little similarity to each other.*

On the other hand, it is difficult to list the clustered risks on separate lines without drawing arrows between and among them. There are three types of linkages in the clusters. First, among the biologic risks there are the risks of multi-problem children, those who are deaf, mentally retarded, and epileptic for example. These children tend to have chromosomal damage (from rubella, for example) and previously would probably have died. Second, there are the biologic risks of high-risk pregnancies, from mothers too young, too old, or with other such dangerous conditions. And finally there are those linkages between environmental/social risks with biologic risks in a two-way relationship.

* I do not mean to imply that the conditions here are never part of larger syndromes or associated with clustering risks; certainly sickle cell anemia, for example, is usually associated with poverty and virtually always with blacks. But blackness and poverty (the core source of the clustering) are only circumstantially linked, while sickle cell anemia is genetically linked.

The clustered risks are not really different from the independent ones in this regard. Most of the independent conditions could be depicted as the result of or the nexus of a cluster of biologic and environmental events. Cancer, for example, is often related to levels of carcinogens in food and air. Because this is not a medical paper, I am not going into such etiological and epidemiological richness. These linkages are either too clear (e.g., the spread of typhoid from unsafe water) or too contested (e.g., the role of childhood cholesterol or salt intake on adult disease) to merit new policy action. The linkages demonstrated by the clustered risks, however, are more interesting because they suggest the complex interaction of health care and other variables with the health of the individual and the sadly prominent reality of the social etiology of childhood disease and poor health. This social confluence, however, does offer hope that many child health problems, because they are not independent risks, may be alleviated by the same health or health-related programs.

In the outline, references will be noted for some of the health risks. Following the outline, I discuss some of the most important problems or clusters of risk. I should note here that I have not completely organized the risks into "critical periods" or temporal sequences. This would require far greater sophistication in pediatrics, and developmental biology, and developmental psychology than I possess. Some clear evidence or critical ages for serious insult is presented in the discussion of specific risks.

Infant Mortality: Its Causes and Correlates

The United States infant mortality rate (deaths under one year/ 1000 live births) was 22.4 in 1967 (the latest figures available). The rate ranged from 16.6 in Utah to a high of 35.5 in Mississippi, with individual areas going above and below those state figures. The discrepancy between rates for whites and non-whites (there is no direct measure of poor vs. non-poor) has been growing in recent years; in 1967, the national rates were 35.4 for non-whites and 19.6 for whites. In 1950, the non-white rate was 66% higher than the white rate; by 1965 it was 85% higher. Overall infant mortality rates are highest for blacks and American Indians. (White House Conference on Children, Profiles of Children, 1970, pp. 48-49; hereafter referred to as WHCC) Put in terms of life chances, a non-white child runs a risk one and one-half times greater than a white child of death in the first month of life and a risk more than two and one-half times greater in the first year. (Huron Institute, 1972, Vol. 1, p. 371)

The leading causes of infant mortality, according to a study by Diana Hunt (1969), include:

- 1) postneonatal death which occurs largely as a result of infectious diseases, particularly influenza, pneumonia, diarrhea, and enteritis;
- 2) prematurity;
- 3) failure to obtain prenatal care;
- 4) failure to identify and care for high risk pregnancies; and
- 5) the total environment.

According to Profiles of Children (WHCC, 1970), more than 50% of infant deaths are accounted for by premature births, congenital malformations,

and postnatal asphxia. (p. 50)

The evidence for the effect of prematurity on survival chances is generally indirect. Because all women do not receive early prenatal care or determine the exact date of conception, it is sometimes hard to determine whether a baby is in fact early, on time, or late. The usual proxy measure is birth weight, with low birth weight (under 2500 grams or 5 lbs. 8 oz.) being highly correlated with prematurity and strongly with risk of death. Some studies (c.f. Werner, Bierman, & French, 1971) have separated low birth weight from actual prematurity and compared their separate and combined effects. They found that there are some normal, full-term infants who are of below average weight; these infants show little residual damage or risk to survival. There are also premature infants, who tend to be under normal weight, but between 1500 and 2500 grams; these infants, if raised in a nurturing environment, catch up with their full-term, full weight peers, although a poor environment or other health risks may prove more damaging to these fragile children. Finally there are the infants of very low birth weight, usually very premature and often with other problems; these are the infants who experience great risk of death in the early weeks and months of life and who frequently demonstrate irreparable damage in several areas--health, size, mental ability, psychological functioning. This risk is vividly shown by the figures on infant mortality rates by birth weight.

<u>Birth Weight</u>	<u>Infant Mortality Rates/ 1000 Live Births</u>
1000 gms. or less (2 lb. 3 oz. or under)	919
1001 - 1500 gms. (2 lb. 4 oz. - 3 lb. 4 oz.)	548
1501 - 2000 gms. (3 lb. 5 oz. - 4 lb. 6 oz.)	207
2001 - 2500 gms. (4 lb. 7 oz. - 5 lb. 8 oz.)	58
2501 - 3000 gms. (5 lb. 9 oz. - 6 lb. 9 oz.)	19
3001 - 4500 gms. (6 lb. 10 oz. - 9 lb. 14 oz.)	9
4501 gms. or over (9 lb. 15 oz. or more)	13

Source: National Center for Health Statistics (1960), Department of Health, Education, and Welfare (As cited in WHCC, pp. 41 and 90)

There is a significant difference in the epidemiology of the neonatal and postneonatal and infant mortality rates. There is less difference along socioeconomic and racial lines in neonatal rates. Death rates in the first week of life (early neonatal) have declined much more slowly than those in the first year of life over the last forty years. (WHCC, p. 49) The dynamics appear to be thus: the early neonatal mortality curve is asymptotic and is practically at its apparent asymptote in all of the Western countries. It has remained virtually constant in the U.S. since 1950. As more and more of our population becomes urbanized and hence has access to hospitals, this rate (which represents the limit of modern technologic and medical control) has given up as much ground as it is

going to. The late neonatal and postneonatal curves, however, are much more influenced by environmental conditions (their first major reduction, as Dubos (1959) has pointed out, being during the Sanitary Revolution of the 1800's) and socioeconomic status. It is these rates in the United States which account for most of the difference in white and non-white death rates. This is shown most sharply in Indian infant death rates: early neonatal rates are better than the national average, but those for late neonatal and postneonatal (7-28 days and 28 days-1 year) are significantly worse. (AAP, 1971, p. 153) Much of this mortality is due to infectious diseases such as tuberculosis and trachoma, which are no longer real threats to middle class children. In the hospital, care can protect the newborn child from disease; there is no genetic susceptibility or weakness. But when the child is living in the hogans in which everyone including livestock lives and eats on the floor and the standard of living is pitifully low, the impact of health care is much less. A study of a six-year Navajo-Cornell Field Health Project comments:

The effectiveness of contemporary medical technologies is far more dependent on the socioeconomic circumstances of the recipient in the case of the infant than it is in the case of older children...Modern medical technology has relatively little to offer infants who are located in an unprotected home environment. (McDermott, et al., 1972, p. 29)

Rural and semi-rural areas in general report death rates at least 10% above the national average. (Hunt, E., 1970) This is probably due both to environmental conditions (poverty, poor housing, infections, etc.) and also to a residual elevation of the early neonatal rate because of the lack of access to hospitals and of

unattended home deliveries. A combination of largely black, rural, poor populations explains the geographic concentration of excessive death rates in the southern states. Over 50% of all non-white births in 1967 occurred in just 16 southern states and Washington, D.C.; the 15 southern states with above-average rates also have the highest concentration of families with incomes below \$2000 a year. (Huron Institute, 1972, Vol. 1, p. 380)

Childhood Mortality

The risk of death in the first year of life is higher than in any other year until 65. (WHCC, p. 95) Even if the infant rate were equivalent to the early neonatal rate of 15.0 (i.e. if the only deaths occurred in the first week), the risk of death would still be higher than every other year until age 55. But once past this critical year, childhood mortality is an extreme and rare occurrence. The lowest average death rates are for children ages 5-14 (0.4), with the second lowest for ages 1-4 (0.9). (WHCC, p. 95) In characterizing the patterns of these rates, Richmond and Weinberger (1970) say, "In general, the patterns of leading causes of death among non-white children tends to resemble the pattern among white children at least a decade earlier." (p. 27) Leading causes of deaths between ages 1 and 4 are:

- 1) accidents, by a wide margin; rate much higher for non-whites
- 2) congenital malformations
- 3) flu/pneumonia (for non-whites); cancer (for whites)
- 4) cancer (for non-whites); flu/pneumonia (for whites)

5) meningitis (WHCC, p. 64)

Of the accidents, 33% are from motor vehicles and 22% from fires and explosions, both situations over which medical science or health care has virtually no direct control.

Childhood Morbidity

Many of the traditional diseases of childhood--infectious diseases, tuberculosis, pneumonia, gastroenteritis--have been virtually eliminated except for the poor and the non-white, who still do not receive basic immunizations or live in healthy environments. The rates of tuberculosis for non-whites are four to five times higher than for whites at the same ages. After a dramatic drop since 1959, both cases of and deaths from diphtheria have slowly risen since 1965--startling evidence of slippage in basic immunizations and disease control. (All data from WHCC, pp. 26-7; also see Huron Institute, 1972, Vol. 1, p. 411; North, 1967; and Siegal, 1966)

I would like to discuss an important concept about infant and childhood morbidity. The concept, of a continuum of physical and psychological insult from infant death to mental retardation and other non-fatal deficits, was first proposed by Knobloch and Pasamanick in their work on the epidemiology and sequelae of reproductive casualty. (Knobloch & Pasamanick, 1966; Pasamanick and Knobloch, 1960, 1966) Noting the association between pregnancy complications and fetal and neonatal deaths due to brain injury, they hypothesized that among the survivors

...there must remain a fraction so injured who do not die, but depending on the degree and location of trauma, go on to develop a series of disorders extending from cerebral palsy, epilepsy and mental deficiency through all types of behavioral and learning disabilities resulting from lesser degrees of damage. (Pasamanick and Knobloch, 1960, p. 304)

This is probably the single most important hypothesis and organizing concept in child health care today. It is enabling researchers and practitioners to gradually bring together a wide variety of factors, syndromes, illnesses, and environmental conditions into a systematic model of the epidemiology, etiology, and process of morbidity in children. Needless to say it is not completely accepted; nor is it able to deal with all of the problems of child health. It is important as a policy tool primarily because of the overwhelming evidence, some of which has been presented above, that many reproductive complications and causes of infant mortality are associated with social conditions and inequities: poverty, race, lack of prenatal or regular medical care, poor nutrition (of both mother and young child), and unsafe living conditions. As one major review of knowledge about child health and development recently stated:

Although the figures become vaguer and less easily definable, it should be obvious that with a higher level of perinatal mortality, goes a greater degree of morbidity. And, one can safely assume that the long-term effects of perinatal morbidity--though not precisely documented--must be taking a considerable toll in human potential on the survivors. Longitudinal studies of the detrimental effects of prematurity on physical development, psychologic and intellectual growth are reported in the literature. "It appears that the non-white infant is subject to an excessive continuum of risk reflected at its extremes by perinatal, neonatal and infant death, and in the survivors by a reduced functional potential." (Richmond & Weinberger, 1970, p. 29; they quote Birch, 1968)

Studies cited in the Huron Institute report (1972) provide evidence of higher levels of visual and hearing defects among premature infants,

differences on psychological tests between low birth weight and normal children (with greatest impairment at the lowest birth weights); the impairments appeared to be linked to neurological defects.

...in identifying that segment of the general population which is subject to the highest rate of infant loss, we are, in effect, defining a group in whose surviving children we can expect to find not only a high incidence of present poor health, but a higher than average prevalence of primary neurological damage representing the aftermath of excessive exposure to hazards of gestation, birth, and early life. (Birch and Gussow, 1970, p. 13)

One primary example of neurological damage is mental retardation. Between 100,000 and 200,000 mentally retarded babies are born each year. (There are now about 2.5 million retarded children under 20 in all.) Seventy-five per cent of these children are mildly retarded (educable); 15% are moderately handicapped (trainable); 8% severely (many trainable); and 2% are profoundly retarded (unable to care for themselves). (WHCC, p. 51) One fourth of these cases can be linked to genetic abnormalities, infections such as German measles during early pregnancy, birth accidents, or postnatal infections or traumas, events which also are prime causes of infant mortality. Suspected causes for the others include inadequacies of prenatal and perinatal care, nutrition, child rearing, and social and environmental opportunities. (Ibid.) (See also Joint Commission on the Mental Health of Children, 1970; and AAP, 1971, pp. 60-6)

This primary neurological damage is usually confounded by the undue effects which adverse conditions have on premature children. The most serious effect of this continuum of harm has already been

shown in the elevated postneonatal mortality rates of premature infants. One of the most damaging conditions appears to be poverty. The Kauai study (Werner, et al., 1971; Werner, 1967) of children with varying degrees of birth complications concluded that even severe complications (such as prematurity) could be compensated for by a good postnatal environment, but for the poorest children the cumulative effect of birth complications and environmental risks resulted in lower IQ scores. Drillien (1964) reaches similar conclusions, except that he reports noticeable handicap in infants of very low birth weight (under 3 1/2 lbs.) regardless of the family's socioeconomic status or the mother's "competence." Another factor which he also found to affect premature infants (or ones with other kinds of perinatal stress) was family stress.

The chances of this residual damage being promptly discovered, treated, or possibly offset by extremely advantageous conditions of growth in the early years is severely diminished when one considers that prematurity and low birth weight are both more frequent occurrences in non-white and poor families to begin with, with each factor contributing separately to the correlations. (See Huron Institute, Vol. 1, pp. 398-404) (See also Birch & Gussow, 1970; Werner, 1967; Werner, et al., 1971; Vernon, 1967; Drillien, 1961; Glass, et al., 1971, Robinson & Robinson, 1965)

The issue of cognitive deficits as the result of prematurity and low birth weight is a very complex one. I do not have enough knowledge of education, intelligence tests and child development to enter that debate directly. According to Wiener (1962), only one of eighteen studies done since 1940 failed to find premature children lower in

IQ at some point. Others, such as Werner, Bierman, & French (1971) and Drillien (1964) find differences depending on socioeconomic background, family environment (e.g., education of mother, presence of books in the house, stable family), and whether the birth weight was only marginally below normal or extremely low. It would seem that, just as we cannot really determine whether the fact that race is correlated with prematurity rates even when socioeconomic status is controlled for is due to genetic differences or the generational impact of poor health, poor nutrition, and poor living, so too we cannot expect to isolate the separate contributions of prematurity or other birth risks and of later environmental factors on cognitive development and performance until we have a sufficiently large number of high-risk infants who are not raised in high-risk environments. My own suspicion is that much of the correlation will disappear. To suggest why this might be so, I will now examine two sets of causes for low birth weight, prematurity, or death, which will both be shown to be primarily socioeconomic in origin.

Maternal Condition and Risk of Reproductive Complications

The general health, nutritional status, and physical history of the mother are related to the risk of reproductive complications, which are in turn associated with damage to the fetus resulting in death or illness. Birch and Gussow (1970), in their comprehensive review, show that almost every complication of pregnancy and birth (such as toxemia, anemia, prematurity, and postnatal asphyxia) that is potentially damaging to the infant is more prevalent among the poor and the non-white. The Huron Institute report (1972) also summarizes

these findings. Briefly, all of the following conditions of the mother have been linked, to greater or lesser degrees, with chances of reproductive complications, usually measured by the variables of prematurity, low birth weight, or infant death:

- 1) Maternal age. The handful of large-scale studies have consistently shown "that the mortality rate is extremely high for mothers younger than 15, declines to a minimum in either the early or late twenties, and increases fairly sharply thereafter." (Menken, 1972, p. 47) "The increased risk of prematurity may be the most important medical aspect of teenage pregnancy....Whether because of biologic or environmental factors that affect the infant directly or indirectly through prematurity, the infant born to a teenage mother has a much higher risk than infants of somewhat older mothers of suffering specific severe handicaps." (Ibid., pp. 49-50)
- 2) Parity. The risks of pregnancy increase with the number of children or the number of pregnancies a woman has had. Infant mortality tends to increase after the third birth in all but the highest socioeconomic class, where it increases after the fourth birth.
- 3) Spacing of pregnancy. The risks associated with many pregnancies are increased if the births are close together. Huron (1972) cites data showing that "high parity and youthfulness are also associated with a high risk of mental deficiency in the child, with the highest risk mothers being under twenty at third parity." (Vol. 1, p. 386) Again because of their earlier age at first pregnancy and their larger families, low-income mothers are more likely to be in these high risk groups.
- 4) Other conditions. There is some ambiguous evidence that short stature of mothers is associated with low birth weight in infants. (C.f. Birch & Gussow, 1970; Huron, 1972, Vol. 1) There are also the obvious medical risks associated with pregnancy for hypertensive, diabetic, obese, and anemic women, as well as those who, along with their husbands, possess dangerous or harmful gene combinations. The impact of nutrition on reproductive complications is discussed below.

Because large-scale studies with sufficient numbers of women in each category to assure statistical significance are difficult and rare, it is hard to determine the extent to which actual pregnancy

factors or environmental ones are responsible for the ultimate health problems of the child. For example, Menken (1972) believes that "fetal, perinatal, and to a lesser extent, neonatal mortality appear to be caused primarily by factors related to the pregnancy itself, while postneonatal mortality is attributed more often to environmental causes." (p. 47) Likewise, the factors of maternal age and parity can be interpreted biologically to indicate the woman's full physical maturity and capacity to give birth, the declining effectiveness of her reproductive system, or its degree of previous stress. These factors may also be interpreted sociologically to indicate the preparedness of the mother to accept either a first or a fifth or sixth child, the readiness of the society to accept her pregnancy (particularly for young, unmarried girls, but also for women having a child very late in life), and the capacity of the family to accept, nurture, provide for, and love the new child. This dual interpretation of ultimate outcomes seems quite plausible for many of the sequelae of prematurity.

Nutrition and Child Health

The role of nutrition in human health and development and its causal relation to an hypothesized continuum of permanent damage beginning with fetal and infant deaths, is probably the most exciting and prominent topic in early childhood studies. Of primary concern are the impact of the mother's nutrition, both over her lifetime and during pregnancy and nursing, on the potential of the child for survival and optimal growth; the infant's nutrition during

the critical period of brain development and myelination; and the impact of childhood nutrition on growth, health and intellectual achievement. References to some articles are made in the listings above; of the general references, those most concerned with nutrition questions are Birch and Gussow (1970), the leading spokesmen in the field; Gussow (1970); Profiles on Children (WHCC, 1970); and Richmond and Weinberger (1970).

This is a huge topic, one which deeply fascinates and concerns me. I am doing continuing research and writing in the area, and hence it is hard to stop in mid-sentence and summarize. One way to begin is to briefly summarize some critical definitions. Merrill Read (1969) has put them well:

In considering this problem it is important to define malnutrition and to distinguish it from hunger. Malnutrition is viewed as a condition characterized by an intake of one or more nutrients so insufficient that specific clinical signs appear or retardation in physical development is observed. The term severe malnutrition refers to a clinical syndrome arising from long-term protein-calorie malnutrition in early childhood. This is called kwashiorkor and results in extreme stunting of growth accompanied by edema, skin ulcers, and hair discoloration. Specific syndromes of malnutrition may result from consumption of diets inadequate in certain nutrients such as protein, vitamins or minerals; anemia, rickets, and vitamin deficiencies fall into this classification. Demonstrable growth retardation may or may not occur. Hunger, on the other hand, is the state of insufficient food intake required to provide a general sense of well-being and to meet immediate energy needs. If sufficiently prolonged and serious, hunger may develop into malnutrition. The damaging effects of hunger and malnutrition are more devastating in the child because he is a growing organism. Hunger may be easily and immediately relieved with food, whereas malnutrition requires prolonged rehabilitation. (Read, 1969, p. 1)

What I would like to do here is to develop the argument for the importance of nutrition in its most powerful, overarching, and hence most controversial form: the direct link between malnutrition and

mental deficiency. The argument as presented follows that of Bonnie Kaplan (1972) who has done the most comprehensive review and synthesis of current research which I have found. Her work is particularly interesting because she directly confronts the issue of what would constitute "direct evidence" of this linkage, evidence which she contrasts with that from studies which examine related or intervening variables such as prematurity rates, birth weight, and so on.

An adequate model for a research design to answer the question, "What is the effect of malnutrition on mental development?" should strive to fulfill the following criteria:

1. Both the duration of malnutrition and the extent of physiological and psychological effects should be assessed.
2. The approximate period in life during which the malnutrition was experienced should be consistent across subjects.
3. All other physiological variables--prematurity, chronic disease, congenital malformations--should be carefully controlled.
4. Environmental factors--family income, socioeconomic class, diet, parents' education--should be carefully controlled.
5. Mental development should be assessed with a variety of measures: tests for IQ, for cognitive development, for perceptual and motor skills, and perhaps for emotional and personality development.
6. Tests of psychological development should be given periodically over a decade or more, to examine the permanency of the effects. (Kaplan, 1972, pp. 327-8)

She examines several major studies which have attempted to answer the question posed, none of which meets all of the requirements of her model. This, she says, "does not justify ignoring the research. Several of the studies [one of them is Chase and Martin (1970), cited above in the list]...present relatively thorough methodologies and very convincing results....The definitive study to prove causality and to study the significance of timing, environmental compensation, and

so on, has not been done; it cannot be done because of the ethical restrictions on experimentation with human subjects." (*Ibid.*, pp. 328, 330) Kaplan is not bemoaning this fact at all. She proceeds directly to summarize the partial evidence which we do have for a continuum of casualty. Again I quote at length:

Nutrition begins to affect the life of a baby long before it is born. Intergenerational data...demonstrated that the nutrition of the baby's mother, grandmother, and perhaps other ancestors can influence his own chances for unimpeded physical and mental development. One mechanism by which the intergenerational factor expresses its influence appears to be via the mother's stature, particularly the structure of her pelvis....

Another way in which nutrition affects a child prenatally is through intrauterine environment....multiple births,...the mother's consumption of critical nutrients, ...income....correlated with income level. This relationship is easy to understand,...especially since high protein foods are always more expensive than starches and other carbohydrates.

After a baby is born he is still vulnerable to the effects of nutritional deficiencies, some of which are reversible...The timing of the rehabilitation [is] crucial; the earlier the treatment, the greater [is] its effect.

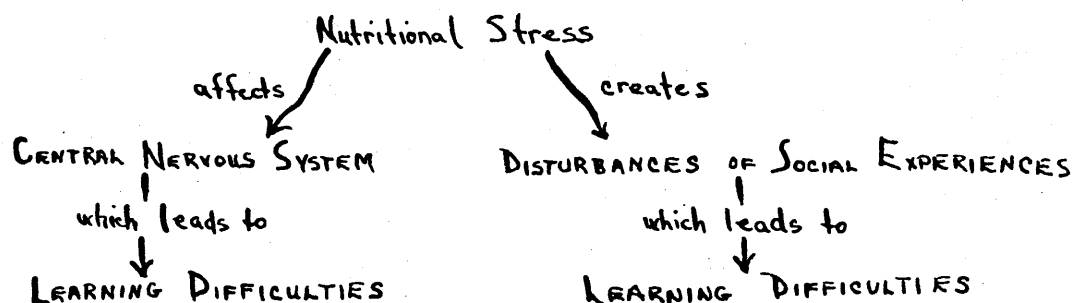
...It is most helpful to imagine a continuum of casualty, to borrow the term used by Knobloch and Pasamanick. Beginning 6 months before birth and lasting until 6 months after birth, cell division is providing rapid growth of brain tissue. Nutritional deprivation during this period can severely retard psychological development, and it is probably irreversible unless treatment begins before cell division ends. From 6 months through the first or second year of life the rate of brain growth is still very rapid, mainly by protein synthesis, but nutritional rehabilitation still can be partially effective. Beyond this point the rate of brain growth is significantly slower, and [one study] found negligible effects of treatment on IQ scores beyond the age of 4 years. (Kaplan, 1972, pp. 330-331)

Very similar conclusions on the effects of malnutrition in childhood have been drawn by Dr. Charles Lowe, research director of the National Institute of Child Health and Human Development, in his testimony before the Senate Select Committee on Nutrition and Related Human Needs, Jan. 28, 1969 (as quoted in Richmond and Weinberger, 1970,

p. 33). Lowe notes an additional barrier aside from ethical constraints, to proving childhood effects from nutrition: that studies of adult populations have been the backbone of the science of nutrition and that the science has therefore failed to recognize "that nutritional needs of the infant are unique and bear only tangential relations to those of the adult." (Ibid.)

Elias (1971) has also summarized evidence on the behavioral consequences of malnutrition in infancy. Despite the problems with the studies which Kaplan has outlined (particularly the omission of data on the onset and duration of malnutrition), she found that there is evidence of major environmental differences, such as family disturbances and housing conditions, between malnourished children and controls from the same social class which suggest that it may not be either poverty or biological insult which is the crucial factor in the linkage. Further, the prolonged hospitalization which may be required in cases of moderate or severe malnutrition may itself have a retarding effect on the child's development.

One model of the effects of nutritional stress suggests how the organic and social factors probably interact to create outcomes attributable to neither along. It has been proposed by Birch and Cravioto (C.f., Birch, 1968) who have both been involved in a long term, and still continuing, research and action project in Guatemala on nutrition and human development. The model is depicted below.



The left hand chain illustrates the type of sequence which Kaplan (1972), Elias (1971), Chase and Martin (1970), Sandstead, et al., (1971), and Read (1969, 1972) are primarily concerned with. The right hand chain illustrates the effects of nutrition on learning through three indirect mechanisms:

- 1) Loss of learning time. Because poorly nourished children are more susceptible to infections and illness, they may miss more days of school.
- 2) Changes in concentration or motivation. The malnourished or hungry child may be distracted from learning by the actual physical pain or at least discomfort of an empty stomach.
- 3) Loss of the social interaction involved in meal times. Particularly for children whose world contains so little cultural and social stimulation, stability, or sharing already, the loss of the social setting of mealtime is not a trivial one. (Kaplan, 1972, p. 331; See also Gussow, 1970; Callahan, 1971; Read, 1969, 1972)

Birch notes that the overall effect is to make the malnourished child a second-class citizen in his ecological environment:

Children who are ill-nourished are reduced in their responsiveness to the environment, distracted by their visceral state, and reduced in their ability to progress and endure in learning conditions....Consequently...there is a reduction in the profit which a child may derive from exposure to opportunities for experience. (Birch, 1968, p. 596)

Elias (1971) reviews one hypothesis which suggests a circular relationship in which the malnourished child is also apathetic and makes fewer demands on his mother and in turn receives less attention, less stimulation and less food, thus becoming more malnourished and apathetic. Hence failure to thrive or maternal neglect, usually considered as psycho-social problems (c.f. Talbot & Howells, 1971), may have an organic root in a child's hunger.

Even if I had more faith than I do in the impact of education on life chances or human development, I would not want to leave the

impression that the only justification for concern about maternal and childhood nutrition is their effect on the learning process or brain development. Certainly if malnutrition can cause permanent brain damage it should be avoided at all costs. The trouble with arguing from this standpoint about American children is not so much that we can't prove the linkage with mental deficiency, but because the levels of malnutrition in this country are generally well above those suspected of causing permanent harm. This is not to say that there are not some American children who in fact do have kwashiorkor, marasmus, or other severe nutritional conditions, particularly in parts of the South, in urban ghettos, and on Indian reservations. (Citizens' Board of Inquiry, 1968; Kotz, 1969; AAP, 1971)

To return to the world of policy for a moment, there is at least one serious program implication of this current focus on nutrition as a part of a complex syndrome of critical insult to child development. If one does not accept the reasonableness of the evidence outlined above as sufficient for program action, basing policy on scientific child development grounds may hamstring policies arguable for other purposes. As Birch and Gussow have forcefully put it:

It would be misleading, if one were to conclude...that malnutrition and illness were significant hazards to children only in early infancy or only as a consequence of their "permanent" effects. It must not be overlooked that the child's present hunger and illness also immediately affect his level of attention, his interest, his motivation to learn--in short his achievement in the classroom. Unless we intend to feed children today it may be interesting, but unimportant to their prospects, to decide whether the effects of yesterday's hunger will continue to affect their mental development tomorrow. Since mental development is a process, perhaps only relatively more vulnerable to interruption at one point than at another, it is difficult to imagine that anything in the environment which interfered for a significant time with learning could fail to

affect mental development. The real children in our classrooms are not like animals in the laboratory who can be rehabilitated at times appropriate to their development and convenient to a researcher. In the real world it may actually be quite unimportant whether the effects of nutritional stress are organic, biochemical or emotional. Whether or not damage is permanent must be irrelevant in the face of a deprivation which is; the same children whose mothers were ill-fed and unready for pregnancy, who are born into poverty and survive an infancy of hunger and illness are seldom miraculously saved in the third act. (Birch & Gussow, 1970, pp. 262-63)

But there is considerable subclinical malnutrition which can also have serious effects on children. To understand this malnutrition and hunger, one needs to shift focus from outcome measures of mental deficiency to epidemiological measures of nutritional deficits or gaps of various kinds. The argument behind this approach has been well put by Young (1971):

The present state of knowledge does not permit establishing quantitative functional relationships between nutrition and the achievement of outputs in health, education, and other important areas. Thus, for purposes of a preliminary analysis at least, it is expedient to assume that nutritional deficiencies (failure of people to consume required quantities of essential nutrients) are harmful per se, and that a sufficient objective is to attempt to eliminate these deficiencies. (Young, 1971, pp. 8-9)

I would differ with Young only to say that it is more than expedient to make such an assumption. It is equitable and just to make it since the majority of Americans, who are well-fed, do not need to justify their own consumption of food on grounds other than that it is better to eat than to go hungry. Let me now turn, then, to what we know about the nutritional needs of American children on the basis of equity.

To begin, there are a group of studies which compare the nutritional status in different socioeconomic groups (primarily in England).

Kaplan summarizes them as follows:

In virtually all such studies the following relationship seems to hold: Relatively poorer women, when compared to women of higher economic classes, have less nutritious diets and give birth to babies who are more often of low birth weight or perhaps premature, are less healthy, are more likely to have congenital malformations or serious diseases, and have a somewhat lower probability of living beyond a few months. (Kaplan, 1972, p. 323)

Although stress could also be a factor explaining this difference, studies in which dietary supplements have been given to low-income women have had significant effects in producing healthier babies. (Ibid.) Protein deficiencies seem both the most related to income and the most harmful to the fetus. Maternal malnutrition, often over several generations, is an important part of the continuum of damage suggested by Pasamanick and Knobloch and central in the clustered risks which I have listed, even at levels observed in the United States.

It would seem obvious that except in times of brief nutritional deficits (such as the Dutch famine for 18 months during World War II due to deliberate Nazi policy (Stein, et al., 1972) or the shortages in England during the war) it is unlikely that the child of a malnourished mother is himself going to be well-nourished in early childhood. As long as the child is being nursed, he can act as a rather effective parasite on the mother's nutritional reserves and can gain quite adequate nutrition from her milk. But after weaning, his nutritional status may decline precipitously.* Malnu-

*This is possibly one of the environmental reasons why children in large poor families tend to display lower intelligence or greater risks; if the children are closely spaced, the safe time of nursing is cut short when the next sibling comes along.

trition in infancy and early childhood has been shown to be associated with increased susceptibility to infectious diseases and higher rates of mortality from diseases not usually fatal. (Huron Institute, 1972, Vol. 1, p. 408; Scrimshaw, 1965.) Moreover, infections frequently depress an already poor nutritional status by interfering with vitamin synthesis and nitrogen balance (Birch & Gussow, 1970), as well as often just reducing the appetite.

The extent of malnutrition in the United States is not clear. Hunger USA (1968) identified 256 counties in poor regions and in some urban areas as hunger areas; 10 to 15 million people were estimated to live in families eligible for food supplement programs. The Department of Agriculture (1968) has estimated that while only 9% of families with incomes over \$10,000 had what it called "poor" diets, 36% of families with income under \$3000 did. A "poor" diet was one containing less than 2/3 of the recommended allowances of one or more essential nutrients. Sixty-three per cent of the upper income groups had "good" diets; only 37% of the poor did.*

Probably the most widely used measure of the inadequate nutrition is iron deficiency anemia. Fomon (1970) has shown that infants of low birth weight and those born to malnourished mothers are among those most likely to develop iron deficiencies. Hunger USA noted that between 30 and 70 per cent of children in poverty areas suffer

* There is some evidence that the majority of Americans are becoming more poorly nourished; the percentages of "poor" and "fair", as opposed to "good", diets at all income levels had increased since the last survey!

from some degree of iron deficiency anemia. Woolsey (1971) summarizes evidence from several sources on rates of anemia among children of different races, family income, and age. Rates and severity of anemia are higher in very young children (under 2 years) with some improvement even for poor children as they get older. Many people have hypothesized that this is part of the "weaning syndrome" mentioned earlier: that until the child learns to fend for himself a little better in the world, he is at highest risk of insufficient nutrients. (Gutelius, 1969) Huron Institute (1972) also summarizes evidence on a variety of nutrient deficiencies and diseases (rickets, scurvy, vitamin A and protein), all of which affect poor and/or non-white children most severely. The Ten-State Nutrition Survey of 1968-1970 (originally the beginning of a national survey) was concentrated in poor, mainly southern states. It found that among preschool-age children 51% of the boys and 56% of the girls had substandard levels of vitamin A, while 21% of the boys and 27% of the girls had low riboflavin levels. These low levels were income-linked. Looking at vitamin C, they found that of 12-23 mo. old infants, less than 10% of those from families in the highest income quartile had low levels, while 40% from the lowest quartile did. Similarly, for children aged 2-6 less than 4% from high income families were deficient while over 30% from the lowest income homes were. (Cited in WHCC, pp. 65-66.)

Finally, there is evidence of general retarded growth among low-income children which can presumably be linked to inadequate nutrition. Owen and Kram (1969) found that in a sample of Mississippi preschool children the poor were on the average smaller than the more affluent,

and at greater nutritional risk. Among migrant workers' children in upstate New York, 38% were in or below the third percentile for height as compared to norms standard in pediatric practice. (Richmond & Weinberger, 1970, pp. 33-34) These latter authors go on to note a subgroup difference.

In fact as many as 25 per cent of the children in the age group from one to three years were in the third percentile; and 13 per cent of the total were below the third percentile. Those most affected were from two to three years of age; the number below the third percentile was ten times the number predicted.

...Bergstrom [who did the original study] suggests that what is noted may relate primarily to nutritional intake--the young infant is most often provided with his nutrition by his mother, and the older child is better able to fend for himself; whereas at two to three years of age, he is no longer being fed, is being weaned from the bottle, and is not quite able to "compete" for the available food.

...He suggests the possibility that "catch-up growth" occurs between age three and five in these children.
(Ibid., p. 34)

What we do not yet know, from this or other studies, is the extent to which that catch-up growth actually wipes out the deficits of the intervening years. The fact that the older children are still below height norms in unexpected numbers suggests that this recovery is not complete; on the other hand, it is not quite clear how directly height is related to health, IQ, or performance in later life.

I have focused this discussion on the clustered risks associated with malnutrition in young children. There are several other nutritional issues which many consider as important as outright malnutrition or potential developmental damage on the health and welfare of children. First is the controversy over the role of cholesterol and high-fat diets in childhood on adult cholesterol levels, atherosclerosis, and heart disease. Second is the role of high

salt diets on later hypertension and heart disease. And finally there is the growing problem of obesity among all Americans but especially among children.

These problems affect all American children. They are independent risks since many Americans choose to eat high cholesterol, over-processed diets of empty calories or simply to eat too much for the exercise they get. They are also clustered risks to the extent that high-starch, high-fat diets are cheaper and easier to prepare with limited facilities than balanced, higher protein ones. These are risks which have just recently begun to attract research interest and money; we do not yet know very much at all about their causes or seriousness. It is likely that they will remain with us after the basically social problem of malnutrition has at last been dealt with.

Concluding Remarks

Many of the risks listed at the beginning of this section have not been explicated or even defined. Space does not permit me to examine such vital problems as dental health (whose most significant aspect Richmond and Weinberger (1970, p. 34) term "the sheer magnitude of the problem"; 25% of all children ages 5-14 have never visited a dentist); child abuse and neglect; accidents; other environmental risks (lead poisoning from paint or automobile exhaust; rat bites, substandard housing, radiation); and the whole range of mental health problems. The gaps in references on the list of risks do not correspond to the relative volume of research in each of

those areas, but rather to the uneven coverage which I have given to them in my research. This is not in way of an apology, but only to reemphasize that this chapter is not meant to be a comprehensive or authoritative picture of child health needs. Rather I hope to have suggested in several of what I do regard as central areas of need the types of knowledge which we have about child health, the implications of that knowledge for policy issues, and the ethical judgments inherent in any interpretation of the data and testimony. The next part of this chapter will provide a broader analytic framework in which to understand those needs which I feel to be the most pressing, those clustered risks of cumulative and often geometrically increasing harm.

Part 2.

Impact of the Health Care System

I discussed the developmental and individual health consequences to the child of both biological and environmental health risks in early childhood, risks which tend to cluster together. A full picture of the dimensions of disadvantage and risk means also looking at the ways in which the health care system itself--its structure, content, and operations--interact with the child and his health.

These systematic effects of health care delivery are particularly relevant to a discussion of public policy options since most large-scale programs (either service or financial) cannot be delicately monitored as to their effects on all individuals, but they can be studied for broader impact (frequently expressed in statistical terms). What we are interested in is how programs (e.g., Medicaid, Maternal

Infant Care) or structures (e.g., private fee-for-service medicine) themselves determine the outcomes of illness in groups of people. Some of the following discussion will stray beyond specific impacts on young children. This reflects a first systematic disadvantage to children: our methods of health care delivery and financing underinvest in children and their health needs. Hence data are frequently not analyzed specifically in terms of children.

According to a special analysis of the federal budget for fiscal 1973, federal health outlays for children will amount to \$2.1 billion in FY 1973, one-half for indigent children (primarily the estimated 11 million to be covered under Medicaid). However, this represents only 10% of the federal health budget, of which 60% goes to those over 65 and the remaining 30% to persons 22-64. (OMB, 1972, p. 169) Even if one accounts for the higher rates of hospitalization and generally poorer health of the elderly, the figures are not nearly comparable. This six-to-one advantage for the elderly is greater than the generally accepted ratio of need (elderly-to-general population) of four-to-one. Looking at the figures from another perspective, according to recent SSA figures, the federal government pays for only 16% of total health care costs for children, but 54% of total costs for the aged. (It pays even less--11%--of costs for those 19-64.) The basic reason for this difference is that national policy has accepted provision of a minimum level of health services as a right for all Medicare beneficiaries. Such a right does not exist for children. (See Table 3.) Since extending similar coverage to children is essentially a question of belief or political expediency, there is no sort of evidence or

Table 3.

Estimated Health Care Expenditures by Population and Income Groups*
(in millions of dollars)

	1973 estimate	% of total
Total, all recipients	19,915	100%
Aged (65 and over)	11,772	59%
Other adults (19-64)	6,042	30%
Children and youths (0-18)	2,100	11%
Indigent, total	8,867	100%
Aged (65 and over)	4,586	52%
Other adults (19-64)	3,224	36%
Children and youths (0-18)	1,057	12%
Nonindigent, total	11,048	100%
Aged (65 and over)	7,187	65%
Other adults (19-64)	2,819	26%
Children and youths (0-18)	1,043	9%

* Does not include foreign nationals receiving health care services outside the United States

Source: U.S. Office of Management and Budget, Special Analyses, Budget of the United States Government, 1973, Washington, D.C.: U.S. Government Print-Office, 1972, p. 169.
(Percentages added)

testimony which would be persuasive as to why it should be done; but certainly there is no evidence that coverage for children would be anywhere near as expensive as coverage for the elderly on a per capita basis.

In order to examine this and other disadvantages, we have broken the system effects of care down into three categories: Access to Care, Appropriateness, and Quality.

Access to Care

Access can be a significant variable in determining the health risks of childhood. It would be part of structural and impact evaluations. Its impact can be divided into five subcategories: geographic, financial, psychological, legal, and group.

Geographic Access

The well-established facts of the uneven distribution of physicians and of health resources in the United States have special implications for children. Seventy-five per cent of pediatric care is given by general practitioners. Pediatricians have remained at about 3-5% of practicing physicians since 1949 (American Academy of Pediatrics, 1971). In addition, pediatrics is still an almost regional specialty: one-third of all pediatricians practice in New York, Massachusetts, and Pennsylvania. Since the increase in the number of physicians in general has kept pace with the population, pediatricians themselves are not in shorter relative supply today, but due to the decline in the number of general practitioners, primary physicians for children are. Primary physicians include pediatricians

and one-fourth of the general practitioners (studies have shown that is the average amount of their time devoted to children). Combining these groups yields a ratio of one primary physician to 2,248 children under 14 years of age in 1969, as opposed to a ratio of 1 to 1,483 in 1949. (AAP, 1971)

As with all specialties, pediatric practice is not attracted to areas of scarce population (a pediatrician needs a much larger population base to yield a profitable practice than does a general practitioner) and/or economic deprivation.

The causes [of maldistribution] are not peculiar to pediatrics, but the problems are intensified because, as a general rule, the geographic areas with the highest proportion of children are also the most remote or most unattractive or most poverty stricken. (AAP, 1971, p. 216)

To take a particularly extreme example: 47% of the American Indian and Alaskan native population (a group for whom the government has assumed full responsibility for care from the private sector) is under 15 years of age (vs. a national figure of 31%). Yet in 1969 there were only 20 pediatricians among the 390 physicians in the Indian Health Service. (AAP, 1971, pp. 152-6)

Within urban areas, poor children, like their parents, may be isolated from private practitioners and can obtain clinic care only via relatively expensive and time-consuming public transportation routes. Very rarely are such networks designed to facilitate access to health facilities by their clients.

A new kind of geographic access problem has sprung up with increasing government intervention in the delivery system. If a child is poor or otherwise defined as at risk, his access to Maternal and Infant Care and Children and Youth Programs is dependent on whether a large array of external forces have moved

so as to locate an operating program in his locality. Although all the states receive money for maternal and child health and crippled children programs under the formula grants of Title V of the Social Security Act, the usefulness of those programs for the individual child depends on how his state has implemented the program: whether it has put the money into service delivery instead of basic research or manpower training and whether the service programs are in his area.

Flouridation, a public health measure which has consistently reduced dental caries among children by about 60% where it has been introduced into public water supplies, is again a geographic variable beyond a family's control, usually being dependent on a community both large enough to have a public water supply (thus denying its cheap benefits to rural children) and politically willing to enact enabling ordinances.

Financial Access

The issue of financial access is basically one of equity. Our society values health and accordingly treats health care as a valuable, scarce good. It also treats it as basically a private good. In general medicine is practiced privately, hospitals are controlled privately (whether non-profit or for profit), and care is financed by private insurance (although government's contribution continues to rise) or direct payment by the private individual.

Thus access often becomes a function of the ability to pay for services. Thirty to 40 million persons in the country do not have access to adequate health care because of inadequate income (Somers,

1971, p. 63) As the costs of health care have skyrocketed, some form of health insurance has become almost essential to permit sufficient and necessary utilization of services without severe economic hardship. In 1968, about 80% of the civilian non-institutionalized population had some amount of hospital insurance (coverage of ordinary hospital charges for inpatient care); 78% had surgical insurance and 65% had medical insurance (for inpatient physicians' services). These percentages tend to overstate actual coverage, which may vary from total service benefit provision to a very limited indemnity, with deductibles, co-payments, and other limits. In 1968, only 36.0% of total consumer health expenditures were covered by health insurance (up from 27.7% in 1960). Furthermore, the gap in coverage is largest for those who can least afford to meet their medical expenses. While only 8% of the under-65 population with family income of \$10,000 or more do not have hospital insurance, 64% of those with family income under \$3,000 lack such basic protection--and as noted above, this is the most prevalent type of insurance. (All of these figures are from HEW.) This relationship holds true for women covered for obstetrical services and hospitalization. (White House Conference on Children, 1970, p. 32)

These figures refer to the entire population. While specific breakdowns for children are not available, the structure of our health insurance system suggests some special disadvantages for children. Insurance in general covers less than 40% of total consumer health expenditures; the largest gap in this coverage is for preventive, ambulatory care (routine office visits not related to a specific illness, non-hospitalized diagnostic tests, outpatient

drugs, dental care of all kinds, and psychiatric care). The American Academy of Pediatrics estimates that 90% of child health care is delivered in an outpatient setting and that about one-half is for health maintenance, not care for acute illness. Hence current payment mechanisms stressing inpatient coverage mean that "payment for quality comprehensive child care is grossly inadequate." (AAP, 1971, p. 233) In view of this, the recent extension of compulsory Medicaid coverage to include diagnostic screening services for children under 21 is extremely important, although the regulations are not yet being enforced.

One of the greatest failures of child health programs to this date has been their inability to provide real alternatives to welfare medicine for the poor or disadvantaged. There may be a few new Maternal and Infant Care centers and so on, but for the most part what is available is still indigent medical care in which, as Yerby noted, "patients are forced to barter their dignity for their health." (Yerby, 1966; see also Kosa, et al., 1969; de Vise, 1960; Duff & Hollingshead, 1968) Breslow (1969) has summed up the issue well:

Perhaps more important than the relatively small total amount which the federal government allots to child health is the disproportionate amount which goes to forms of care that are increasingly unacceptable to the American people and to the medical profession. Less than one fourth of the present federal funds for child health go to maternal and child health and OEO programs that are being designed to overcome older patterns of second-rate care based on the "means test." (p. 329)

By far the largest program providing funds for child health services is Medicaid, a state-controlled program which, unlike Medicare, is a part of the welfare system. Medicaid pays for care

for about 1.5 million children. An analysis done for the OCD Child Health Care Policy Task Force (Minnesota Research Systems, Inc., 1972, p. 7) reports the following estimates of Federal Medicaid expenditures for children (i.e., not including state matching funds):

FY 71	\$0.7 billion
FY 72	\$1.0 billion
FY 73	\$1.3 billion

Looking back at Table 3, two things are evident. First, the bulk of all services to poor children are clearly coming through Medicaid. Second, one of the estimates is off since Medicaid appears to be paying out more than the Federal totals in Table 3. Since the figures in the Task Force document are unpublished ones, we would tend to consider them high. (They may also include pro-rated administrative overhead and non-service payments, which the OMB figures do not.)

In summary, certain children are at special risk simply because their families are poor in a society of fee-for-service, privately insured health care which has not yet fully resolved those inequities. Furthermore, such inequities in some areas place children at special risk--because the reimbursement system has not been designed to meet their special needs.

Psychological Access

The general literature on the indignities and inconveniences associated with being poor and seeking medical care is extensive and persuasive (e.g., Duff and Hollingshead, 1968; Strauss, 1967) Psychological barriers may be interpreted in at least two ways. One theory, along the lines of a "culture of poverty" argument, suggests

that the poor or the disadvantaged place a low priority on health care, are less aware of illness in themselves and their children, and hence utilize services less and in less appropriate ways (breaking appointments, waiting until the last minute, using emergency rooms, etc.). The problems inherent in using utilization data to draw conclusions on preferred behavior in a system of pervasive financial and geographic inequities in access are still not recognized by some health analysts. I have previously discussed this problem which can be accurately summed up here in a statement by the American Academy of Pediatrics:

There is little evidence of general apathy or of different attitudes or expectations toward the importance of health care, and most or all of the observed variations in utilization can be explained in terms of barriers imposed by the institutional arrangements for health care of lower socioeconomic groups. (AAP, 1971, p. 28)

The second interpretation of psychological barriers, and the one which I am most impressed by, is that medical care systems are often not designed or operated in ways which make them attractive, acceptable, and open to certain people. Required means tests before delivery of care, refusal of many non-profit private hospitals to admit (even in emergency rooms) indigent patients, lack of bilingual personnel in non-English speaking areas, hours of service which are not convenient for the users, and staff and space inadequate to avoid crowding and excessive waiting--all of these circumstances are going to have significant impact on the willingness or ability of poor people to use medical care until it is absolutely necessary.

Such circumstances may arise from a variety of causes: racial prejudice, lack of sensitive training for personnel, inadequate resources

of health institutions. One of the most important causes is the lack of fit between the system's goals and attitudes and those of its clientele. The emergency rooms and outpatient departments of large teaching hospitals are the major source of ambulatory care for many disadvantaged children and their families. While these institutions treat most people who come in the door, their main function is not a service one, but an educational and research one. The system's goals are maintaining a steady and adequate flow of "teaching material"; there are rarely staff arrangements facilitating continuity of care by one physician over time. These are clearly not the mothers' goals; "psychological distance" is one way of describing that misfit in expectations. Another important misfit is that between a welfare mentality which is primarily concerned with keeping chiselers off the rolls and helping only the "deserving poor," and the demand for service by a family requiring medical care through no fault of its own.

Evidence exists that changes in the systems involved can result in vast improvements in psychological access. The evidence from neighborhood health centers, the Cornell Medical School Medicaid experiment in New York, and the prepaid group practices of H.I.P. and Kaiser-Portland in covering poverty groups supports the notion that patterns of usage are to a great extent influenced by the ways in which care is given. (Sparer and Johnson, 1971; Goodrich, Olendzki, and Reader, 1970; Bellin & Geiger, 1968; Colombo, Saward, and Greenlick, 1969) The evidence suggests that outreach efforts, new types of personnel with less professional distance (e.g., pediatric nurse practitioners vs. pediatric residents), and strong

consumer/community participation in health care systems are important variables in improving psychological access. (Notkin and Notkin, 1970; Silver, Ford, and Day, 1968; Patterson, et al., 1969)

Legal Access

Health care is generally not a legal right in this country. Certain groups directly under the "protection" of the federal or state and local governments (Indians, armed forces, and institutionalized persons) usually have some legislated provisions for health care. With the exception of the armed forces, these "rights" are often unmet and untested. (See for example AAP, Lengthening Shadows, 1971, pp. 152-156, for brief summaries of levels of care for Indian and institutionalized children.)

For the rest of the population, legal access to payment or reimbursement for health care is increasingly available through employee-employer or individual contracts for health insurance, through federal entitlement (Medicare), state entitlement (Medicaid, workmen's compensation, institutions for the mentally ill, etc.). Within the private sector, entitlement to continuance of care by a physician, once initiated, and protection against incompetent care are provided by civil law (and criminal in the case of some malpractice suits) under medical ethics statutes. However, at this time, there is no legal right to services per se. For example, a county cannot be required to provide a hospital unless it decides to do so; the last doctor in a rural area cannot be prevented from leaving even if his departure means no medical care for the residents. Recently, the federal government has taken a major step by proposing

new regulations which would require all hospitals and nursing homes receiving money under the Hill-Burton program to provide a specific amount of free care for the poor, at the risk of otherwise losing federal funds. Other recent suits are attempting to force private non-profit voluntary hospitals to open up their doors to the poor or lose their tax-exempt status as charitable institutions.

Special Groups

Much of the evidence previously cited on developmental and environmental criteria has emphasized the effect of race on health risk, health status, and utilization of health services, independent of income, education, or geography. This gap in access, in other words, seems to be in certain ways identifiable by group membership. Haggerty (1970a) charges that one of the most important weaknesses of the present health care system is its inadequate care for the disadvantaged. He includes among the disadvantaged: "The black, the American Indian, the poor, the rural, the poorly educated, the agricultural migrant, the physically and emotionally handicapped, and high-risk groups, such as children born out of wedlock"; as well as the growing number of suburban middle-class children without medical care. (p. 78) The American Academy of Pediatrics (1971) has also reviewed evidence that membership in these groups is statistically a predictor of not only suboptimal health status but also access to health care. The Academy also includes another special risk group: children not living with their families (i.e. children in foster care, in institutions for the mentally retarded or disturbed, or in various detention centers).

I have suggested the importance of prenatal care, especially for high-risk pregnancies.

...on purely obstetric grounds, certain groups of the population should receive the highest available levels of maternity care. There are primigravidae, ...grand multiparae, women with a poor obstetric history, young mothers with rapidly recurring pregnancies, women of short stature and poor physique and health, mothers having illegitimate pregnancies, and the socially underprivileged... (Birch and Gussow, 1970, p. 156)

Yet several studies have shown that poor mothers, those most likely as a group to be at risk, are the least likely to receive adequate prenatal care or indeed any at all. (Lesser, 1964)

The implications of such gaps can be seen in terms both of special risk and of equity. Migrant workers clearly live lives of much greater risk than most of us do: of accidents, epidemic diseases (gastritis, schistosomiasis, etc.), malnutrition, tuberculosis, parasitic infestation; yet they receive even less health care. A 1966 study (Siegal) found no migrant children with complete immunizations; the AAP in 1970 estimated perhaps only one-third had any immunizations and one-sixth had reasonably complete ones. (AAP, 1971) In 1969, after six years-of the Migrant Health Act, two-thirds of the 900 counties where migrants live temporarily still had no grant-assisted programs. And combining all sources of funds for those programs, the average per capita expenditures were \$12 annually. Special efforts to meet the special needs of these groups are not adequate. But in addition, migrant workers are frequently denied equal access to more general care programs. In most states, they are ineligible for state general assistance (which carries with it eligibility for some minimum or even ample amount of medical care) or Medicaid (via welfare eligibility) because they do not meet residency requirements.

Dr. E. S. Rabeau, Director of the Indian Health Service, has pointed out the dimensions of a problem which applies not only to the Indian population but also to migrant workers and to a large portion of the black and rural poor populations. His remarks emphasize that these groups live under unique conditions which will require special, broadly defined health measures to assure them of even token equity in achieving healthy lives.

[There is] a continuing need for added emphasis to achieve acceptable health levels for Indians and Alaskan natives and comparability to the general population....The inferior health status...is largely due to lack of basic sanitary facilities, unsafe water supplies, gross unsanitary practices, poor and crowded housing, inadequate nutrition, emotional problems inherent in a transitional culture, impoverished socioeconomic status, and limited education. (Quoted in AAP, 1971, p. 156)

In other words, the healing effects of even the "Sanitary Revolution" of the 19th Century which Dubos (1959) described as the underlying cause of our good health has not yet reached all Americans.

Appropriateness of Care

Access is fundamentally an economic function, having to do with the supply and distribution of scarce resources. Even certain kinds of psychological access can be improved merely by economic changes (increasing the staff or floor space of facilities, hiring more bilingual personnel), although full access is certainly non-economic. Appropriateness, as we shall use the term, is a non-economic, qualitative function, having to do with the content of those resources, the fit between what medical science and the health care system find appropriate practice and the needs of young chil-

dren.* We have identified three subcategories of appropriateness which have a major impact on children.

Medical Science

One way to evaluate appropriateness is to ask: are medical efforts being directed at the most critical risks for children. Those risks include poor prenatal environment, low birth weight, infant mortality and morbidity, sensory and neurological deficits (especially if undetected early), child abuse, accidents, dental problems, and malnutrition.

The large gaps which presently exist in even basic prenatal and maternity care for all mothers and for detection of sensory and neurological deficits seem to stem not from a lack of medical interest but from problems of access, of providing the benefits of medical science to all mothers and children in ways they can afford and use easily. On the other hand, for the abused or neglected child, the child with certain congenital problems (such as sickle-cell anemia), lead paint poisoning, nutritive diseases or deficits, or with multiple handicaps, there is frequently little in the medical or health system of use to him. Pediatric medicine has not taken some of these critical areas as its own first priorities; nor have the wider systems which also have impact on the child. Recently, the problems of sickle-cell anemia and lead paint poisoning have begun to be dealt with more realistically and adequately. There is also growing concern about child abuse and redefining the boundaries

* See the earlier discussions in Chapter 3 of how appropriateness (also called "accountability," "scope and importance," etc.) has been incorporated into some evaluation designs. Two of the best real-life examples are Geomet (1971) and NAS (1972).

of it and other social illnesses of childhood. However, the other areas remain inappropriately unexplored.

No theory has been developed to predict the effects of changes in nutrient intake on the overall performance of an individual....The lack of a scientifically established relationship...makes a definitive nutrition adequacy standard an unattainable goal at the present time. (Segal, 1970, pp. 18-19)

Perhaps the most important deficiency is lack of a clear-cut national commitment and policy that care of the handicapped child is an integral part of total responsibility of the government to its population. Unless the problem of the handicapped child is viewed as a humanitarian obligation--as a concrete, important domestic problem, requiring adequate economic and technical support--the creative programs required will not be forthcoming. (As quoted in AAP, 1971, p. 58)

In other words, the system of health care does not always direct its resources in ways that reflect the rates of incidence, prevalence, and severity of childhood risks. Thus some children may end up being at greater health disadvantage than others not because they are sicker or poorer but because resources have not been devoted to their illness or problem. We are talking not about rare childhood diseases, but common ones.

Manpower

While the American Academy of Pediatrics estimates that 90% of child health care is ambulatory and 50% of that care is for health maintenance (AAP, 1971), the allocation of pediatric resources often reflects a more acute illness, inpatient, physician care pattern. We do not suggest that this threatens the child when care is in fact given (although the evidence strongly suggests that American children undergo unnecessary tonsillectomies and adenoidectomies; Perrott and Chase, 1968). But, going back for a moment

to an economic analysis, money spent building unnecessary hospitals, or hiring only doctors instead of mixing them with pediatric nurse practitioners, at some point will mean that a child (probably a poor, high-risk one) will not have free or inexpensive outpatient care, screening facilities, or immunizations available. In those circumstances, the consequences of his illnesses or handicaps will be more serious than if the system had been organized differently.

In addition the specialty of pediatrics is not yet really willing or able to serve all children.

Pediatrics remains a predominantly urban, white, middle-class and upper-class phenomenon. The poor, non-white and rural dwellers have had to utilize other resources. (AAP, 1971, p. 100)

The number of medical students who choose pediatrics as a specialty is, in our system, an uncontrollable variable, since there are no national quotas or even recommendations for such distribution. Since pediatrics is one of the lowest paid specialties, this aspect of our free enterprise medical care system probably works against children; it is no accident that while there are not enough pediatricians in the U.S., we have twice as many neurosurgeons per capita than other western countries.

Appropriate manpower may also mean improving psychological access by increasing the numbers and responsibilities of community workers in health facilities. Of particular importance to children might be workers who could assist mothers in taking care of their own and their children's nutritional needs and serve as advocates to obtain safer environmental conditions if necessary.

Delivery Modes

We have touched on this broad area by noting the misfit between the ratio of preventive to acute care services which children need and the actual ratio of such services. Several other aspects are important:

- a) Although what developmental evidence we have suggests the critical nature of the first four years of life, the health (and educational) systems do not facilitate systematic access to children below school age. Particularly in the case of chronic physical disease and handicapping conditions, early detection and treatment may be critical to enable the child to develop to his full potential. (AAP, 1971, pp. 56-7) Achieving this early access has been one of the major goals of Head Start health programs.
- b) Because of their isolation in the Department of Agriculture, the federal food supplementation and school lunch programs are not oriented either toward delivering the most possible food to hungry children and families or to best identifying such people and their actual needs. Furthermore, it is difficult to use food programs either as trade-offs against other programs (e.g., treatment of nutritional diseases) or as a part of a comprehensive community health program, because HEW lacks jurisdiction over them.
- c) In general, the present health care delivery system is disorganized and frequently cannot deliver comprehensive, coordinated care to children and their parents. Some of the relevant studies of comprehensive or family-care programs will be discussed briefly in Part 3 of this chapter.

It might be useful, however, to note here some findings from a major review of evaluations of prepaid group practice which relate specifically to the needs of children and mothers. The review cites strong evidence that prepaid group practice:

- tends to increase utilization of preventive health services (general check-ups, prenatal and postnatal care), especially for usually underprivileged segments of the membership (non-white and Puerto Rican);
- tends to increase readiness (and reduce delay) in seeking care;

- tends to reduce substantially the disparity between high and low socioeconomic groups in the use of services;
- tends to increase the use of specialists for children and for childbirth; and
- clearly leads to decreased premature birth rates and rates of perinatal mortality (standardized to age of mother), for both white and non-white populations. A slight lessening of the gap between white and non-white rates was also reported. (Donabedian, 1969)

The importance of the findings on specialist care is suggested by results from a comparative evaluation of OEO Neighborhood Health Centers and other providers in which the performance of solo general practitioners in obstetric and pediatric cases was clearly inferior and "woefully inadequate". (Morehead, Donaldson and Seravalli, 1970, pp. 9-11)

Difference in the quality of care, when added to the problems of finding one's way through our complex, pluralistic system, suggests that the burden of private medicine, increasing specialization, and multiple public programs weighs most on those families whose children are frequently most in need of comprehensive care.

This fragmentation (of preventive and restorative services and institutions) imposes a much larger coordinating burden on parents in poverty than is imposed on better situated parents. (AAP, 1971, p. 29)

Quality of Care

The issues of whether outpatient department care is really worse than the services of a private physician, whether pediatricians give better care than non-specialists, of whether "charity medicine" on the wards of teaching hospitals or county hospitals is dangerous as well as demeaning, and of to what extent a good doctor-patient

relationship (implying both continuity and mutual respect) improves one's chances of cure or comfort--all these issues are both of critical importance and without clear answers. In fact, there are few analytically useful definitions of the terms used in the questions. Although the work of Duff and Hollingshead (1968), Strauss (1967), Kosa et al. (1969), and Goodrich et al. (1970), among others, has looked seriously at the issues of an inequitable distribution of quality of care, with rather dismaying findings, the implications for policy are not easy to draw. There is no evidence, however, that children suffer any more risks from poor quality care than the population at large, and hence I will leave the issue unresolved.

Part 3.

Major Federal Health Services Programs for Children

To succinctly and systematically characterize current health programs for children proves to be a remarkably difficult task. A recent study of federal programs for disadvantaged children lists 65 whose activities in some way involve health and/or nutrition. (Huron Institute, 1972; this listing owes much to the prior work of the Appalachian Regional Commission, 1970.) These programs run from the multi-billion dollar Medicaid program (Title XIX of the Social Security Act) which covers indigent and medically indigent persons of all ages in every state, to categorical grants for demonstration Children and Youth projects (59 in all), to dental health research fellowships not directly targeted at children, but potentially valuable to them.

Of the 65 programs, approximately 24 are service programs: seven for special medical problems (immunizations, family planning); eight for comprehensive medical care for special groups such as Indians, migrants, mothers and infants; seven nutritional or feeding programs; and two multiservice programs of which health is one component. (Actually one of these "programs" includes all Office of Child Development efforts: Head Start and Parent-Child Centers, both multiservice programs, and Health Start, a comprehensive medical care experiment.) Two of the 65 programs (Medicaid and Aid to the Blind) provide financial support to children. Of the remaining programs, 16 are for research, 11 for training (of both service workers and professionals), five for technical assistance, and seven for other functions (construction of facilities, information dissemination).

Two of the service programs, Maternal and Child Health Services and Crippled Children's Services are block grants to the states to be used at the latter's discretion for services, research, training, and facilities. Virtually none of these programs, as far we have been able to determine, have been evaluated or monitored in ways approaching the model presented in Chapter 3. Several major evaluations are presently under way, but findings have not yet been published. The dearth of evaluations is not fully explained by the weaknesses of current evaluation tools; there is little doubt that better studies could be carried out given the methodologies and constraints of today.

The purpose of this thesis is not to describe and document the activities of these programs. But most of the major programs initiated in the 60's with explicit "arguments" promised the application of knowledge to the needs of more children (primarily poor ones) and the extension of knowledge by documenting the positive effects of social service programs on the health of mothers and children. The extent to which these promises have been or are being fulfilled should provide some indications of both the rightness of the arguments and the limits of our ability to determine that rightness. Hence, some of the major programs are discussed here, from the standpoint of what they have proven about the impact of health services on the health of disadvantaged children. I have not examined more indirect programs, such as manpower training, research on child health, various scholarship and fellowship monies, or technical support. These represent 39 of the 65 programs, or 60%, although they account for a far smaller percentage of funds appropriated.

Maternal and Infant Care Program

This program, funded under Title V of the Social Security Act, operates 60 projects across the country. It was established in 1965 and had \$42.7 million appropriated in FY 72. As paraphrased by Weckworth, its broad legislative mandate sets its goals as the "reduction of incidence of mental retardation and other handicapping conditions caused by complications associated with child-bearing, reduction of infant and maternal mortality, particularly in areas with concentrations of low-income families without access to pre-natal care, infant care and family planning services." (Minnesota Research Systems, Inc., 1972, p. 13) We do not know to what extent the program has lived up to those goals. One large-scale evaluation of the subgoal of reduction of infant and maternal mortality rates, done by the University of Maryland, has completed the data collection phase and is in the process of analysis and release of final results. Preliminary information indicates that the projects are having the anticipated effect on mortality rates. (HEW, 1970)

In testimony before the Senate Appropriations Committee in 1971, the following indirect evidence was presented showing the drop in infant mortality rates in selected cities with large Maternal and Infant Care projects.

Infant Mortality Rate/1000 Live Births

	Calendar Year		% Decrease
	1966	1968	
National Rate	23.7	21.8	8.2
Major M & IC Projects:			
Miami	23.7	21.5	9.3
New York City	24.9	23.1	7.2
Detroit	28.0	26.8	4.3
San Juan	37.6	28.6	23.9

Source: Hearings before the Senate Committee on Appropriations. Department of Labor and HEW appropriations; 92nd Congress, 1st session, p. 3356.

After presenting similar data for metropolitan sub-areas in which drops in infant mortality rates relative to city-wide or comparable statistics followed establishment of large maternal and infant care and family planning programs, Arthur J. Lesser, head of the Maternal and Child Health Service, noted that

while the reasons for the dramatic change...have not been established in a cause and effect relationship, the only new contributory factors that have been identified are the rapid increase in family planning among the poor and comprehensive maternal programs focused on the most vulnerable population. (Lesser, 1969, p. 895)

According to the testimony of Dr. Vernon E. Wilson, HSMHA Administrator, on July 19, 1971, before the Senate Appropriations Committee, Maternal and Infant Care centers have helped reduce the national infant mortality rate from 24.7 per 1,000 live births in 1965 to an estimated 1970 rate of 19.8, a reduction of 20% or four times the rate of decrease of the previous ten years.

As Haggerty (1972) notes in presenting these same data, this was a period in which "several special programs were introduced to deliver better maternal and child health care," and that one can only interpret the data as showing that "on the national scale, there is evidence that input of medical care plus all other factors do affect infant mortality." (p. 107) He raises a methodological issue which suggests why more precise evidence may not be forthcoming even with better studies:

One of the problems of proving the benefits of overall care is the large population needed. For example, to show a 7 per cent change in child mortality (a statistically significant change) one would have to study 6.25 million children. (Haggerty, 1972, p. 107)

Eleanor Hunt (1970) used another method to demonstrate the inverse relationship between infant mortality rates and M & IC projects. She

groups states by the patterns of their infant mortality rate decreases before and after 1965 and then compared that ranking with the number (and average years of operation) of M & IC projects in each group. Her positive findings still do not, of course, control for possible simultaneous effects of Medicaid, other new health programs, or indeed OEO in general on those rates.

M & IC projects were also expected to lead to a reduction of mental retardation and other handicaps. These goals are more complex and require long-range longitudinal studies of cohorts of patients, offspring and controls, since reliable tests for retardation and estimates of the permanence of some handicaps are not applicable until the children are older. Furthermore, such studies require an examination of the entire process and organization of the projects and the milieu in which they are located, since the possibility of exogenous variables affecting apparent rates of increase or decrease is very large and causal linkages extremely difficult to establish. To our knowledge, no such studies are underway or being contemplated. At the present time, the projects are so few and so varied in their environments (some being completely independent and free-standing, and others existing almost only as a legal funding source for certain patients in a community comprehensive care project) that such a comprehensive evaluation would be able to derive few universal conclusions about effectiveness or optimal organization.

Concerning the final goals of targeting the projects on underserved, primarily poor groups, we do have some data which are not too favorable. The Child Health Care Policy Task Force found that only 33, or 55% of the 60 Maternal and Infant Care projects are in

the states with the highest infant mortality rates (Child Health Care Policy Task Force, 1972, p. 11) The Task Force identified 42 counties as being high-risk by using a measure of excess infant deaths (400 or more deaths over the U.S. average; this measure excludes sparsely populated rural districts). Of these 42 counties, 21 had no Maternal and Infant Care project (11 had no federal health care project at all). Furthermore, the coverage of the projects has remained stable for the last several years at around 129,000 mothers and 43,000 infants per year (many of the women receiving only family planning services).

While some projects are surely located in cities or specific areas which are high-risk even though the state does not fall below the median rate, this is still not a very good record for what is, if one uses the proxy measure of infant and maternal mortality rates for risk, the easiest objective to put into programmatic terms. Part of the failure is undoubtedly political: the art of grant-writing and grant-awarding has never been known as a fully equitable, even-handed matter; this is true in virtually any categorical grant-in-aid program. Part of the failure, however, is due to the workings of the medical care system and illustrates one of the difficulties with federal intervention in child health. Merely announcing that money is available for new programs, and even going so far as saying that the money can only be obtained in certain listed areas, does not guarantee that programs consistent with the legislative intent will be set up. If an area is without health resources, there is no one to sponsor the project; if the local medical establishment opposes such "socialized medicine" it can prevent funds from coming in. We point this out not because it is unique to Maternal and Infant Care centers, or even

always true for them,*but to remind the reader of another of the differences between health and education programs: lack of control and reduced ability to target categorical, compensatory programs when they are being imposed not on a universal framework of reasonably similar proportions (i.e., public education systems) but on a vastly heterogeneous, private market with pockets of public responsibility.

Children and Youth Projects

This program, begun at the same time as Maternal and Infant Care, is designed to provide comprehensive care centers for children age 0 to 18 who are poor or medically underserved. (Close, 1969) There are presently 59 such projects with a FY 1972 appropriation of \$47.4 million. The program has been operating at a fairly stable rate for the last few years, serving 464,000 children, but was due for partial defunding in 1973 as projects were supposed to become self-supporting. Its future is presently unclear.

Children and Youth is probably developing the strongest data base of any existing federal child health program. An extensive reporting system has been developed by Vernon Weckworth of Minnesota Systems Developmental Project. (Weckworth, 1971; de Geyndt, 1969) The model, discussed in some detail in Chapter 3, attempts to plot the progress (or regression) of individuals through the Children and Youth system with a goal of having every eligible child at a level of health maintenance. There are masses of data being generated on a quarterly report basis: number of centers with lead paint screening, number

* In fact, although there are no data on the issue, we would suspect the opposite to be the case--maternal and infant care is probably the least political or objectionable of any health program.

with community boards, number and types of personnel trained, number of new registrants, and so forth. The analytic uses to which this data will be put remain to be seen. The present form of the annual summaries is too raw to be useful as a basis for broad policy considerations, although they are useful, apparently, for project management evaluation of outcome.

Children and Youth Projects potentially cover the complete range of critical child health needs because of their eligibility base. It is interesting, therefore, to note the age distribution of the registered children (Lesser, 1969, p. 897):

less than 1 year	41%
1 to 9 years	57%
10 to 18 years	2%

It may seem surprising that so many of the children are infants. This would suggest that expanded Maternal and Infant programs could handle 40% of the present cases. This is important because in many ways Children and Youth makes little sense as a separate program. Although such projects are aimed at comprehensive, interprofessional care and are trying many innovative programs ("Children & Youth Projects," 1971; American Academy of Pediatrics, 1971), they are in opposition to several theories of health care. They isolate the poor into a separate care system; they isolate the child from his family; and they are so few in number that they set up new access inequities by not always being located in areas of the highest need or by not being able to handle all who need care. (Child Health Care Policy Task Force, 1972) And, as I have already said about patterns of child health, aside from the early infancy years, there is nothing in the nature of most

critical health needs of children that clearly requires that they be treated in a separate care system. That is, the organizational mode for a child's health services might ideally be his inclusion in a comprehensive health care delivery system.

Early Screening and Treatment

Screening is a loosely used term. In the strictest sense, it is differentiated from a physical examination or diagnostic examination because it (a) consists mainly of routinized or even automated testing procedures which can be performed by non-professional health personnel (blood tests, vision tests, hearing tests, chest X-rays, etc.); (b) is frequently designed to test for only one or a few conditions of special concern (lead paint poisoning, sickle cell anemia, etc.); and (c) must usually be supplemented by referral to other medical services if abnormalities are detected.

The most popular version of screening now advocated for adults is "automated" multiphasic screening. This format, which incorporates a battery of tests of physical conditions and frequently includes a psychological questionnaire to help locate functional disease or stress, is seen as a modern triage system to separate the heterogeneous masses of people seeking care into those who are truly sick, the "worried well," and so forth. (Garfield, 1970)

For children, however, screening has been advocated for two slightly different purposes. First, it has been used in Head Start and other such programs to provide access for children who would not otherwise have been brought for medical care. The screening should find any undetected diseases, handicaps, and conditions (such as untreated dental caries) which will hinder the child's development if left

unattended. Frequently, however, such programs are operated without the close linkages to follow-up treatment typical of adult screening programs.

Screening has also been advocated for young children as a combined physical-behavioral operation. It is this type of screening which has the greatest implications for child development. Information on the child is gathered on physical, mental, and behavioral factors and combined into an individualized diagnosis and suggested treatment pattern. The importance of such screening is that early detection of handicaps, either physical or mental, can significantly reduce their later severity in many cases. For example, one of the first recommendations of the Joint Commission on the Mental Health of Children, following family planning and systematic prenatal care, was comprehensive pediatric and mental health services for children under three. The Commission was very concerned about the current tendency of children to "disappear" medically between departure from the hospital after birth and school or preschool entrance. These are critical years of rapid development; many conditions (such as metabolic disorders) undetectable at birth become recognizable and treatable.

Corrective measures for children with hearing, motor, speech, and visual handicaps can often prevent interference with learning which leads to retardation in mental development and which may be complicated by emotional problems. Estimates indicate that about 20 to 30 percent of chronic handicapping conditions of childhood and later life could be prevented by comprehensive health care to age five, and approximately 60 percent if health care were extended to age fifteen. (Joint Commission, 1970, p. 33; emphasis added)

The first step in such care must be an early screening and diagnostic effort.

The Kauai study (Werner, Bierman, & French, 1971) found that diagnoses of serious handicaps (physical, mental, or both) at age 2 were generally confirmed by reexamination at age 10. (Evaluations were done by a special team which also used information previously gathered about the children.) Although the study team found that the physicians of Kauai had done rather well in detecting defects generally considered to be recognizable in newborn infants (29 children), and in getting the children into special treatment before age 2 (37 children), the special examinations done by the team picked up 75 additional children with handicaps requiring special diagnostic and treatment services. (p. 44) Furthermore, in the area of mental retardation, although most severely retarded children who also had physical defects had been recognized by family physicians, the two-year special examinations were responsible for first recognizing the mentally retarded children without other handicaps. (p. 45) Even slightly discounting this last finding due to the ambiguity of IQ test scores at age 2, it is clear that a concerted effort to screen all children at age 2 had a substantial payoff in detecting handicaps. The findings are even more important when one considers the fact that Kauai is a stable community, with maternal and health care for all classes substantially more accessible than that in inner city areas or many rural locales.

In recommending at least one good medical and developmental examination for every child in early childhood, the Kauai investigators note that the

...need for closer cooperation between the various professions attending the birth and care of the child is indicated in order to spot early developmental failures in children suffering from deleterious perinatal conditions and to

provide them with a supportive and stimulating environment to minimize the effects of early damage.

Hospital, birth, and physicians' records contain information about the newborn indicating potential trouble--information that is seldom available to community agencies for utilization in planning with the family for the special needs of high-risk infants. (Werner, Bierman & French, 1971, p. 138)

The value of screening and follow-up on a universal basis, utilizing the registry concept, is already widely accepted in other countries. In Europe, especially the United Kingdom and Sweden, there is a general belief that early case finding must depend on universal screening of developmental progress, followed by detailed assessment of all children who show developmental delay. (American Academy of Pediatrics, 1971, p. 56)

A major maternal and child health program analysis done several years ago (HEW, 1966) estimated that a program of case-finding and treatment (done at several ages--0, 1, 3, 5 or 6, and 9) would prevent or correct 30% of chronic handicapping conditions. The first screening could be performed in the hospital; the last two could be carried out through the schools. However, 1-year olds and 3-year olds are generally unreachably. Head Start and its Health Start offshoot have been attempts to reach at least part of this preschool aged population, with an emphasis on screening. (North, 1967, 1970; Hunter, 1970)

Health Start was begun as an experimental program by OCD in the summer of 1971 with two major components: health services and health education. Services were aimed at low income children under 6; 28 projects were funded at a total level of \$800,000. Screening and testing of all children was to be followed up by appropriate treatment, plus health education; the latter goal was supported by a new Biodynamics health education curriculum package.

An interim analysis of the Health Start program (Nay, Vogt, & Wholey, 1972), begun in the summer of 1971, gives some interesting

figures on screening and examinations. The 28 Health Start projects used a variety of techniques to detect health problems, from multiphasic screening to full physical exams; the eligible children were all under the age of 6 (and over 52% were between 3 and 5). Of a total of 6,432 children for whom data were published, 69% had received medical examinations by December 1971 and of those, 66% required no medical treatment.

In their report, the authors note an unresolved issue in the Health Start program which is a critical one for all "screening and treatment" programs aimed at disadvantaged groups. They ask about the relative priorities among possible goals:

- (1) returning children with health needs to a status of basic health,
- (2) developing ways to ensure continued access to health services for children,
- (3) developing ways of coordinating HEW health service and delivery programs.

It would appear that the particular value of screening (or some sort of general examination or testing) at the present time relates to goal (2). Goal (1) requires a full-scale comprehensive health program, while goal (3) is more of a regional administrative matter. Without clarification of these goals, evaluators cannot be expected to draw conclusions about the effectiveness of the program.

But clear goals might not have helped. The funding cycle was so short that few children were kept in the program long enough to expect much effect. Furthermore, a key source of funding for services, the new Medicaid regulations making payment for diagnosis and needed treatments mandatory for children under 21, has never been

effectively implemented.

Family Health Programs

Concurrently with the move to provide special programs directed toward critical health problems or high-risk groups, there has been a trend in federal health programming toward comprehensive health care delivery systems which not only include all ages and all conditions but also all income levels. This trend is usually accompanied by a call for family-centered care and frequently for team practice or the use of new health personnel.

I will not enter here the debate over Health Maintenance Organizations (HMO's), prepaid group practice, medical foundations, or area-wide management networks. Several federal programs serving large numbers of children fall within these categories. (Madison, 1969) They include the OEO Neighborhood Health Centers, comprehensive centers funded under section 314(e) of the Public Health Service Act, some Model Cities health programs, and the new program replacing OEO and 314 (e), Family Health Centers.

The evidence from these programs, many of which are among the most thoroughly evaluated in the health field, strongly indicates the impact that they are having: on changing the health status of users, and particularly of non-white users (Geomet, 1971; Donabedian, 1969; Sparer and Johnson, 1971; Shapiro, et al., 1960); on the acceptability of medical care delivery to the users (Bellin and Geiger, 1968; Elinson and Herr, 1970; Geomet, 1971; Colombo, et al., 1969); and on the quality of care being delivered (Donabedian, 1969; Morehead, et al., 1970). (See also Greenberg & Rodberg, 1971; Peterson,

1971; Saward, et al., 1968; National Advisory Commission on Health Manpower, 1967.)

Intellectually, this movement toward family-centered health care which incorporates both poor and non-poor into one system is quite solid. There is very good evidence from non-federal programs that family medical care is more effective in treating each family member because of the increased knowledge available to providers. (Silver, 1963; Beloff and Weinerman, 1967; Beloff and Willet, 1968; Alpert, et al., 1970) These studies also report high levels of satisfaction by users, both poor and non-poor. (Freidson, 1961; Goodrich, et al., 1970; Bellin and Geiger, 1968; Alpert, et al., 1970)

There are several arguments against comprehensive family care that should be dealt with here. Gordis and Markowitz (1971) attempt to show in a controlled experiment that comprehensive and continuous pediatric care has no effect on health and on medical care utilization. They measured effect by looking at the completeness of immunization, utilization of medical resources, morbidity and mortality and, in a separate study, compliance with drug-giving recommendations. They found no significant differences after one year between a control group and those children and mothers who were offered a comprehensive program provided by a team of professionals. However, their original groups were chosen (randomly) from primiparous adolescents (under 18) who had come to the hospital (Sinai Hospital of Baltimore) for prenatal care. Virtually all of these girls were non-white, and most were unmarried and poor. They are precisely the group most unlikely to receive prenatal care. The level of motivation of all the girls in the sample (who did seek this care) was probably quite high, and hence they could be expected to continue to seek well-baby or illness

care for their infants even if they had no single, convenient source of care. In effect, the controls would seek the services provided to the experimentals. In the second study, the children studied (77) were receiving oral penicillin for rheumatic fever. They had been under treatment for at least one year before the experiment, as compliance tests were made during that time. It seems quite possible that parents, having established compliance patterns for a long period of time, are not likely to radically shift behavior due to a program not specifically directed at that behavior.

This possibility is supported by evidence from Alpert and his co-workers (Alpert et al., 1968; Alpert et al., 1970) Based on a three year study of users of an experimental comprehensive pediatric care center in a Boston pediatric medical center, they found significant differences between users and two control groups on measures of satisfaction, utilization patterns, and planned response to selected medical problems. However, they found no changes in general health attitudes or in planned responses for adult problems. This specificity of attitude and behavior change impressed the investigators: "The fact that the changes are selective and rather specific suggests that the range of services provided will determine the range of attitudes affected." (Alpert et al., 1970, p. 505)

This abbreviated discussion suggests the tremendous impact which family-centered comprehensive care could have if made generally available. It seems particularly relevant to the needs of children since so many of their needs are in fact the outcome of family living patterns or family resource deficits.

Nutritional Programs

Of the 65 federal health and nutrition programs compiled by Huron Institute, 12 are purely nutritional or food distribution programs (including 3 non-service ones), and 8 more have nutritional components (provision of meals, service or education) among their secondary or optional mandates. None of the primary programs are administered by HEW; except for an OEO-CAP Emergency Food and Medical Services program, they are run by the Department of Agriculture. (See Table 4)

We cannot fully explore the workings of all of these nutritional programs. Many major programs are not designed primarily as nutritional programs at all, but rather focus on food surplus control and price stability; the USDA "has not traditionally had any general welfare objectives that go beyond the agricultural community." (Segal, 1970, p. 70) Since evaluation is a process of comparing stated objectives with results, it is not surprising that few federal studies of nutrition programs deal with social or health objectives.

The largest programs by far are not targeted directly at children-- Direct Food Distribution and Food Stamps. There is no breakdown on what portion of the funding for these programs serves children.

The National School Lunch Program is the largest nutritional program directed specifically at children. A school lunch program, however, has no direct impact on those children shown to be at greatest risk of damage from undernutrition, those aged 0 to 2. It could have an indirect impact by freeing family resources to purchase more food for the younger siblings of school age children, a dubious and "filter-down" process. This is a serious stumbling block for which no programmatic answers have yet been supplied. The Special Milk program,

Table 4.

Federal Nutrition Programs Affecting Children

<u>Family and Child Feeding Programs</u>	1973 Expend. (est.) Federal Share (millions of \$)	%Estimated in need being served
Food Stamps (USDA)	2,500.0	{ 49% 58% to
Direct Distribution (USDA)	323.3	
Nutritional Supplements (USDA) (for mothers and children)	16.1	8%
Emergency Food & Medical Services (OEO)	30.0	N.A.
General School Lunch (USDA)	274.7	N.A.
Assistance for Free or Reduced Price Lunches (USDA)	587.7	84%
School Breakfast (USDA)	52.5	20-34%
NSLP Equipment Assistance (USDA)	16.1	N.A.
Special Food Service Programs (USDA): day care centers (year-round) recreation programs (summer)	} 74.0	24%
		27-46%
Special Milk Program (USDA)	97.1	N.A.

Programs with Secondary or Optional Nutritional Components

Elementary and Secondary Education Act (Office of Education, HEW)
(under Title I, Migrant Children, Handicapped Children sections)

School Nutrition Demonstration Projects (OE, HEW)

Head Start, Health Start (OCD, HEW)

Follow Through (OE, HEW)

Indian Children (Interior Department)

Model Cities (HUD)

Indian Health (HSMHA, HEW)

Migrant Health (HSMHA, HEW)

Maternal & Infant Care Projects (HSMHA, HEW)

Maternal & Child Health Services (formula grants : HSMHA)

Children & Youth Projects (HSMHA)

Source: Huron Institute. Federal Programs for Young Children,
Vol. 2 (1972), p. 304, Table 11.1; Citizens' Board of Inquiry.
Hunger U.S.A. Revisited (1972), pp. 10, 12, 49-52.

plus the spotty nutritional components of maternal and infant care projects and state maternal and child health services, are an attempt to provide perinatal supplementation; but the coverage provided by these programs is too uneven to be nationally or even regionally effective. It would seem that food stamps or other indirect routes are the only realistic way to reach these children at present.

Another concern with nutrition programs, especially lunch programs, is that the simple eligibility standard of income may be an insufficient proxy for a more logically sound, but more pragmatically complicated, nutritional status standard. One study has found that between 25% and 42% of children with nutritional deficits (as measured by hematocrit levels and/or weight and height) were being denied free lunches in low income schools because they fell above the income cut-off lines.

By ignoring simple medical indices in favor of family income or impressionistic considerations by school officials, many nutritionally indigent youngsters are excluded from participating in the school feeding program. (Paige, 1971a, p. 261)

The policy issue is whether free lunches are primarily an equity program for poor children (justifying an income scale), a nutritional/health program (justifying a physiological scale), or a combination effort.*

Estimates of need in nutritional terms are very difficult. (Segal, 1970; Elias, 1971) The "recommended daily allowance" figures are non-specific in their derivation (i.e., they are usually set well above laboratory testing levels of damage; Woolsey, 1971). Knowledge of

* It is interesting to note that this same dilemma has emerged in the M & IC program, where an original attempt to reach mothers and infants "at risk" both because of low income per se and because of medical high-risk conditions has evolved into a simpler definition of "high-risk" as "poor".

this inaccuracy tends to conservatively bias policy-making, since it is difficult to justify the standards rigorously. It would appear at this point that the new Health and Nutrition Examination Survey (HANES) now being completed by the National Center for Health Statistics (due to report preliminary findings in December, 1972) should generate valuable baseline data on the nutritional status of the population, particularly those served by federal programs. Participation in such programs will be studied; comparisons of nutritional status between participants and non-participants will be possible. These data will hopefully provide more analytically sound directions for Federal nutrition policy, although, since they are based on the effects of the current fragmented system, they will still not provide much evidence supporting a shift of nutrition and food supplementation programs to HEW or into more comprehensive programs. Although we have not examined the methodological procedures of the study and the following may be a moot point because of better design, we do raise a caveat about an under-estimation of need, despite the margin of error in the standards, because of survey methods.

It would appear, then, that the really poor children, like their pregnant mothers, are not included in nutrition studies or "national samples" because they are not seen by doctors, because their families move about, or because their mothers either do not supply reliable data or do not return questionnaires at all. (Birch and Gussow, 1970, p. 23)

Another problem with surveys and full assessment of nutritional need is the extremely political nature of the inquiries. Hunger USA Revisted (Citizens' Board of Inquiry, 1972) documents in astonishing detail the furor raised by the testimony of Dr. A.E. Schaefer, director

of the Ten-State Nutrition Survey before the Senate Select Committee on Nutrition and Related Human Needs--a furor which resulted in the muzzling of Schaeffer, his ouster as director, the removal of the Survey to the Center for Disease Control in Atlanta, and the cancellation of the planned expansion of the Survey. Similar episodes are documented by Nick Kotz (1969) in Let Them Eat Promises: The Politics of Hunger in America. The greatest concern on the part of those administering and funding the programs seems to be to avoid any acknowledgement of real hunger in America.

In uncovering the thinking that had tailored all the government's food programs, Senators Kennedy and Clark discovered, as others had before, that these programs were designed and controlled by people who never even considered how their clients were effected. Most of all, the Department of Agriculture bureaucrats never considered the rights of the poor, their personal desires, or their possible knowledge about the problems of poverty. The School Lunch Program, for example, received no Congressional examination for fifteen years except for appropriations subcommittees whose influential members flatly opposed more food aid for poor school children. Agriculture Department food administrator Rodney Leonard admitted (after he left office) that department bureaucrats never felt challenged to think about the programs in terms of service to children. (Kotz, 1969, p. 63)

The School Feeding Programs

One of the questions which could have been asked of nutrition programs more frequently is their impact on learning. Evidence concerning the potential brain damage and permanent developmental consequences of maternal or early infant malnutrition has been surveyed above. (Elias, 1971; Birch & Gussow, 1970; Kaplan, 1972; Read, 1969, 1972) It is highly suggestive but not conclusive. Extending the argument, Birch and Gussow (1970) go to great lengths to develop linkages between perinatal stress, subsequent malnutrition, and learn-

ing disabilities. On one level, they would agree with Charles Lowe of the National Institute of Child Health and Human Development (Lowe, 1971) that, "There is a sound scientific basis for the hypothesis that the relief of malnutrition is a key element in modifying the incidence of mental subnormality." (p. 651) They go further to suggest, with others (Garvue, et al., 1971), that functional learning behavior (paying attention, being alert, even attending school regularly), may be much influenced by a child's hunger or under-nutrition. The only nutrition programs of which I have been able to find any pertinent studies are the large school feeding programs.

It is difficult to keep these issues in focus when studying federal food programs. Since the latter have rarely been studied at all, it is relatively easy to hypothesize all sorts of potential effects. In fact, however, there are several reasons why feeding programs will not bring about overnight changes in school performance or health.

...Environmental equalization must be viewed as a longer term process, stretching across two or three generations at least, and we must not expect to overcome within a single lifetime the entire consequences of 15 generations of sub-optimal conditions of life. (Birch & Gussow, 1970, pp. 268-9)

They stress (as Young (1971) also does) the need to separate nutritional goals at least partially from developmental, educational or wider social ones, and to argue the case for programmatic action without incontrovertible scientific evidence or an expectation of quick results:

The health of children and their nutritional status can be immediately improved. It does not matter whether such action produces an immediate improvement in educational performance, since the likelihood that it will hurt performance is zero. (Ibid., p. 272)

The National School Lunch Program (NSLP) reimburses public and non-profit private schools for up to 25% of the cost of student lunches of authorized quality; up to 15 cents per lunch additional is available to help defray the cost of free or reduced cost meals for indigent students. Because of this low matching ratio, free meals are often difficult for local school districts to finance. In fact, recently only 10.9% of school lunches were sold at reduced prices or provided free. (Segal, 1970, p. 93) The School Breakfast Program, although much smaller in budget and coverage, operates in a similar fashion.

A recent analysis of the National School Lunch Program (Young, 1971) contains data which make one cautious about the potential impact of current food programs. Young found that the percentage of all poor children participating in NSLP in 1968 (21.1%) was less than the percentage of non-poverty ones (25.6%). In other words, the program is not even reaching the neediest children--those to whom the nutritional value actually makes a difference--to the same, rather meager extent to which it is reaching the non-poor, for whom it usually only replaces calories and nutrients which would otherwise be obtained from a home-prepared box lunch. Determination of need is left to local schools; there is no nationally uniform test or scale. (Segal, 1970, p. 93)

Young goes on to use rather simplified but comprehensive models to analyze NSLP in the general context of national child nutritional needs--i.e., the extent to which NSLP supplies total child-RDA days (RDA is the Recommended Dietary Allowance, as set by the Food and Nutrition Board of the National Research Council, National Academy of Science.). He develops various models for program operation, varying

the degree of participation and of free lunch provision. He finds clearly that the present NSLP is not even the optimal federal program, given the current allocation level. Providing meals free to all poor children in NSLP schools would result in increasing the number of RDA's provided at less net cost. If, he concludes, one assumes unmet nutritional needs are worthy of being filled per se, then

In terms of output per dollar it appears that the most effective program would be one requiring 100% participation of all schools and free meals for all poor children. This is contrasted with the present program which reaches about 33% of school-age children, of whom 13% receive free or reduced-price meals, i.e., only 2.4 million of the nation's 8.6 million poor school-age children are served. (Young, 1971, p. i)

Young's findings are reinforced in an earlier work by Kahn (1965), who recommended providing free hot lunches to all children. He cites the Swedish experience, where the administrative complexity and undesirable stigma found in a selective program led to a change to a universal meal system.

There are limits, however, to what one can expect even from a more comprehensive lunch program. First, as Young neatly diagrams it, if one looks at the total RDA's per child per year (one day's RDA X 365 X number of children), the NSLP provides only 4% of that total. Fifty-two per cent of the RDA's are allotted to non-school days; two-thirds of the remaining RDA's (= 32% of total) are not provided by NSLP because they represent breakfast and dinner requirements. Even if all children in school received NSLP, this would account for only 16% of their total annual RDA; a school breakfast program could increase that percentage up to a maximum of 32%.

There is some additional evidence that school meal programs can consistently improve the diets of pupils aged ten to thirteen

(Christakis, 1968); one could expect similar results for younger school-age children also. They also may improve attendance rates and pupil-teacher relations. (Garvue, et al., 1971) Paige (1971b), studying anthropometric and biochemical (hematocrit) changes in undernourished children given school lunch versus a control group, also undernourished, found no significant improvement. He suggests however that the one-third of the RDA represented by lunch is not a sufficient critical mass to overcome deficiencies; the length of the test may have also been too short. Additional evidence and testimony are reviewed by Woolsey (1971); Elias (1971); and The Food Research and Action Center (1972).

To regard a Head Start or school lunch program as though it were going to modify the nutritional circumstance of the family is to be self-deluded. The nutritional circumstances of the family can be changed only by a fundamental social modification of food practices, coupled with food availability and understanding by the community. (Birch, 1970, p. 879)

While we share Birch's sense of the ultimate need for broad social reform in order to alter nutritional circumstances, we would only note that part of that long-run improvement begins with each day that a child is better fed in some way than he was before. Joan Gussow has made the point eloquently.

There is, of course, one last argument which can be advanced against the notion that the schools should provide for children's physical needs--that is that we can't prove it will help them learn. The success reported by a few are, in scientific terms, merely anecdotal assertions that good health does make a difference. The hard scientific evidence to support the notion that children's present biological condition correlates with their learning is best described as fragile. There are a few studies--one fairly old one linking blood levels of Vitamin C to IQ, two more recent ones evaluating the effect of iron-deficiency anemia on various measures of functioning. For what it is worth, they have all tended to show that

children who were better nourished did better. But the fact remains that there are, so far as I have been able to determine, no controlled studies which show whether the child who is very hungry is unable to work as well in school as one who is not hungry--or even whether he is just unwilling to. We have no convincing evidence to sustain the notion that too little sleep or too many cavities actually hinder the learning process. That is to say we cannot at the moment prove, on the basis of scientific evidence, that children would profit educationally if they were provided with all the health and nutrition services which they needed.

This being the case, there is one question we need to ask. Is it rational (or humane) to attempt to teach anything at all to a child who is hungry or sick or tired or all three? If it is not, then we shall probably have to go ahead and feed hungry children breakfast, and put sick or exhausted children to bed, and take care of the eyes and ears and teeth and other broken parts of children whose health has been neglected, just because it seems like the only sensible thing to do, and not because we are convinced beyond a shadow of a doubt that such an approach will result in a 23.6% improvement in their SAT scores. At the very least we will then be teaching these children--and their parents--that someone cares about their hunger, their fatigue and their illness--which is much more positive than some of the things they are learning now. (Gussow, 1970, p. 12)

One type of study which is urgently needed is a comparison between direct food programs of any sort and income maintenance as to their relative effectiveness in altering food purchasing patterns and nutritional status. There is a trade-off which needs to be measured between the universal coverage and freedom of choice which an income program would provide (e.g., reaching children aged 0 - 5 without requiring some new institution or intervention program) versus the directedness and immediacy of food programs.

Efficiency and philosophy support the idea that in-kind programs are an inappropriate method of helping low-income households to improve their living standard...But because economic opportunity programs are not capable of immediately attacking the problem of inadequate diet, and because the poor may lack the motivation or knowledge to obtain a fully adequate diet, either with or without more income, it could be argued that even though food programs restrict freedom of choice, they are

necessary as stop-gap measures to save young children from permanent physical and mental damage and to provide older children and adults with diets which give them the best chance of remaining healthy, alert people. The dangers of malnutrition may justify the temporary restriction of choice as a means to insure the recipients are healthy enough to take advantage of programs aimed at permanently improving their opportunities to choose. (Segal, 1970, p. 73)

Medicaid and Financing Plans

Medicaid, by far the largest federal contributor to child health in terms of dollars spent, is a financing mechanism, not a service one. It also covers many people other than children, and data are not readily available pertaining to young children. Medicaid served around 9,000,000 children in 1972. Some of these may have received complete thorough care; others may have been covered only for an emergency room visit. Medicaid cannot guarantee coverage or quality of care; furthermore, the scope of the program, in terms of services covered and eligibility strictness, varies considerably from state to state. A full discussion of Medicaid belongs in an analysis of financing systems and insurance programs. The Child Health Care Policy Task Force (1972) has reviewed the pros and cons of Medicaid and proposed national health insurance schemes. I have noted that health insurance coverage, which presumably would be of most benefit to those without resources to meet medical emergencies, is in fact a function of family income; those most in need are most likely to be unprotected. And although the evidence is not completely consistent, there is every indication that the poor are also ill more seriously and more often.

Thus simple equity would seem to require that any federal subsidized insurance should provide benefit coverage at least as liberal as that of

the middle-class. On this score, it is disheartening to read the Task Force's discussion of the benefits proposed under the Administration's health insurance proposals: the Family Health Insurance Plan (FHIP) and the National Health Insurance Standards Act (NHISA)--the former primarily for indigent families and the latter for employed persons and their families. Neither FHIP nor NHISA (the more liberal of the two plans in its provisions) propose to cover the following services which would seem vitally important to achieve comprehensive care for children:

- psychological or psychiatric services
- nutrition services
- speech therapy
- vision and hearing screenings as needed
- eyeglasses
- hearing aids
- prosthetic devices
- physician nurse and allied health personnel services
- supportive social services
- outreach services.

In addition, FHIP does not cover the following items covered by NHISA:

- prenatal care for and birth of first child
- catastrophic health insurance protection

The Task Force estimates that it would cost only \$50 million to make those two NHISA benefits available to the child population of FHIP, a group most at risk of perinatal complications, lack of access to prenatal care, and congenital problems (often requiring huge medical investments). The group makes no estimate of the cost to include any or all of the other needed services. By comparison, the "welfare" medicine of Medicaid seems comprehensive and liberal. Similarly, existing programs under Title V (Maternal and Infant Care, Child and Youth, etc.) and Section 314(e) have much greater potential for comprehensiveness despite their relatively few numbers of users.

Other Federal Health Programs

Maternal and Child Health Services, Crippled Childrens Services. These two programs, which amounted to a total of \$121,000,000 of federal formula grants to states in FY 1972, are clearly of vital importance in providing health services to children. They have only minimal guidelines, and are state-controlled programs for which no reliable national data exist on coverage, eligibility requirements, program content, or outcomes. An attempt to develop a uniform data base and analysis of these programs is underway by a group at George Washington University. Such lack of accountability is difficult to justify when every health dollar is being claimed by numerous programs. We are particularly concerned that some, if not most, of the funds (at least in MCHS) might not be better spent in expanding the coverage of programs with more demonstrable effects on child health--such as Maternal and Infant Care projects or Neighborhood Health Centers.

Dental programs have been omitted due to the pressures of time, not because we consider them unimportant. When only half of all children under 15 have even visited a dentist, the level of need is enormous, and present levels of federal support are woefully inadequate; in FY 1972 comprehensive dental care projects were expected to reach only 15,000 children.

Indian Health and Migrant Health programs have also not been discussed. Again, the lack of evaluation studies, or even comprehensive descriptive ones, hampers program analysis. In this case, the overwhelming impression one receives of these programs is their inadequacy in the face of tremendous need (American Academy of Pediatrics, 1971; Adair & Deuschle, 1970). What is needed first is sufficient funding to reach enough children in a concerted way to make outcome evaluation meaningful.

Chapter 5.

Conclusion

Health of mind and body is so fundamental to the good life that if we believe that men have any personal rights at all as human beings, they have an absolute moral right to such a measure of good health that society and society alone is able to give them.

Aristotle

I went down into the underworld of London with an attitude of mind which I may best liken to an explorer. I was open to be convinced by the evidence of my eyes, rather than by the teachings of those who had not seen, or by the words of those who had seen and gone before. Further, I took with me certain simple criteria with which to measure the life of the under-world. That which made for more life, for physical and spiritual health, was good; that which made for less life, which hurt, and dwarfed, and distorted life, was bad.

Jack London,
The People of the Abyss
(Quoted in Stansky and
Abrahams, The Unknown
Orwell, pp. 231-32)

There are millions of Americans living in hidden places whose faces and names we never know. But I have seen the children starving in Mississippi, idling their lives away in the urban ghetto, living without hope amid the despair of the Indian reservations. These conditions will change, those children will live, only if we dissent.

Robert F. Kennedy
(Quoted in Kotz, Let Them
Eat Promises, p. 18)

It must be remembered that evidence is never complete, that knowledge of truth is always partial and that to seek certainty is to await eternity

J. Bowlby
(Quoted in Talbot & Howell,
1971, p. 1)

Any plant growing in the wrong place is a weed. For some plants there are very few "right" places, and they are weeds to all of the people all of the time. But even the most beautiful flower can become a weed if it grows outside the boundaries customary to it. Evaluation is not a beautiful flower by anybody's standards, but neither is it the planning equivalent of crabgrass, always out of place and beloved of no one. Much serious and honest belief is placed in the necessity and value of systematically analyzing and measuring our societal efforts at social improvement. Sincerity and conscientiousness, however, are no guarantees of beneficence or rightness. What I have attempted to do in this thesis is explore the boundaries surrounding evaluation and the related policy tools of needs and knowledge which demark the flowers from the weeds. In particular I have sought to outline the extent to which evaluation of child health care programs is a worthwhile and ethically sound enterprise. Three sets of boundaries appear most important: boundaries of health care, boundaries of evaluation itself, and boundaries of scientific knowing.

The boundaries of health care are the limits of what we can effectively hope to do to meet recognized health needs through medical intervention. One determinant of these boundaries is our collective definition of health. If we define health as the absence of disease or certain conditions, then there may be precious little we can do if chronic conditions or diseases are our major health problems. If, on the other hand, we define health as a level of functioning and

expand health care to encompass rehabilitative, care-taking, and preventive services, then health care can be expected to make much greater improvements in health.

Health care is also limited by our technical and biomedical capacity to detect needs and to devise methods (be they curative, preventive, or maintaining) to meet them. Both the discussion of evaluation infrastructure in Chapter 2 and the presentation of critical child health problems in Chapter 4 are relevant here. Some of the limits of health care for children presently most in question are those concerning handicapping conditions (Can we detect them prenatally or early in infancy? Can we prevent them with adequate prenatal care, family planning, etc? What can we realistically expect for the growing numbers of multiply-handicapped children?); nutritional inadequacies (What are dangerous levels and periods of nutritional stress? Can effects be reversed? How does the social environment cause or perpetrate inadequacies? Can we reach young children and mothers even if it is desirable to do so?); and prenatal and other preventive care (How much can we improve the survival chances of babies of poor or high-risk mothers? Can we provide care through new types of personnel, less extensively trained, with less narrowly medical orientations?).

Another general boundary of health care has figured prominently in the preceding discussion: the relationship of underlying social conditions, not controllable by health care, to health status. I have shown that poverty is associated with poor health, poor chances for survival, and limited access to health care. If health care can

remove those relative differences (even if this means greater efforts or "compensatory" programs), then providing the care will improve health and thus remove one of the inequities of poverty; in a sense it will make the poor less poor or unequal. But if the health care only "works" if the patient lives in a sufficiently life-supporting environment (e.g., is above the poverty line), then poverty must be alleviated before health care will show many results. There is growing evidence that the effectiveness of health and nutrition programs alone in the face of otherwise unchanged social environments has been oversold. As Haggerty (1970a) has said, "There is no doubt--the greatest disease of the poor is poverty. (p. 78) In a more discriminating analysis, McDermott and his colleagues (1972) found in studying the impact of technological health care on the children of a Navajo community that

depending on their biologic nature, diseases vary enormously in their susceptibility to medical management in slum conditions....The two conditions that did not require changes in household practices for their control--otitis media and the transfer of tubercle bacilli --were significantly influenced, whereas the two [diarrhea and pneumonia] that did ■■■ require such changes were not. (pp. 28,29)

Lowe (1971) has also suggested that any government strategy to combat malnutrition must include policies of economic action to reduce the poverty which is the root of the problem.

Interest in enlarging the scope of "health care" to enable it to address and affect a wider range of variables is growing. These attempts can be broadly classified as ecological in outlook. Rene Dubos (1959) and others (c.f. McKeown, 1965) have argued that historically the great advances in health indices were primarily

the result of improvements in basic living conditions. Inherent in this outlook is a rejection of the current medical model of single-factor etiology and disease-specific cures. Although some pediatricians (Haggerty, Richmond, Talbot, to name a few) have explored this area, the greatest efforts have been made in mental health.

In the past ten to fifteen years, psychiatrists have begun to recognize the relationship of disease to society. They have begun to move from the concept of single causality to a more comprehensive, ecological approach to mental health. ...toward broader sociobiologic concerns...To operate in the broader system, the psychiatrist must become concerned with developmental processes rather than disease--with preventive mechanisms rather than symptom treatments.... This means that the society itself must be educated so that those human values that promote mental health begin to exert pressure in the decision-making process. (Duhl & Leopold, 1968, pp. 3, 5, 15)

In a provocative article, Bloom (1965) has called for a reconsideration of the long scorned "miasma theory" model of disease, which stressed prevention of diseases and treatment as an arrangement of optimal conditions for the patient to help himself, and did not believe in germ theory or other precise diagnoses. "The major technique for disease prevention was to attempt to modify the environment by removing the sources of the miasma." (p. 336) Bloom presents the theory as having particular relevance to community mental health, but it can be extended to encompass other needs of children. An excess concern with exact etiology and precise diagnosis may prevent medical personnel from designating certain children as "patients" for care or from diagnosing a syndrome which incorporates non-medical factors. Along these lines, Wagner (1970) reports a syndrome found in children resistant to a cognitive stimulation program which included "increased frequency of infections, increased frequency

of child abuse, decreased maternal interest in the child, and marked irregularity of meals provided to the child." (p. 869) The intervention strategies suggested by such a syndrome clearly must address issues that go beyond medical science or the individual child.

We need to know a great deal more than we do about the bio-social interactions that affect children. Chapter 4 has shown trends in both basic knowledge and programs toward more ecological, comprehensive, and interdisciplinary conceptions of children and of child health. Still unexplored are the limits of this ecology. How much of the reality of childhood morbidity is included (ie. how do we handle presently unpredictable genetic conditions or rare diseases which are not part of patterns we can discern)? To what extent will medical practice be able to change? Does the ecological approach suggest new ways of helping children or only new conceptual frameworks for the same events? Do programs so organized show greater effectiveness?

The second level of boundary definition I have dealt with concerns evaluation as a policy tool. Evaluation should be a primary method to illuminate the boundaries of health care suggested above, i.e. it should tell us which interventions work, on which conditions, on what range of outcomes (mortality, functioning, morbidity, learning capacity, etc.), and over what time span. I have discussed at length the degree to which evaluation as presently practiced rarely addresses those questions in a manner broad enough to really guide policy.

There are two types of limits involved. First, the methodological tools of evaluation may not be able to assess ecological or multi-dimensional models of health care in real-life; for example, we can measure infant mortality rates, but not the interaction of nutrition

and maternal stimulation on infant development. Second, the theoretical underpinnings of evaluation may not be adequate to synthesize data into such causal or relational models. In regard to the latter limit, both impact evaluation and systems-oriented models of evaluation seem particularly significant theoretical contributions. In regard to the former, it is my conviction that better evaluations--ethically responsible and methodologically respectable--could be done now, along the lines of the Geomet (1971) and National Academy of Sciences (1972) studies. In other words, we could greatly extend the number and types of evaluation studies of child health programs without creating "weeds". (I am not concerned here with other limits on evaluation such as costs, adequate manpower to carry it out, or administrative cooperation.) I have discussed some of the characteristics of child health needs and measures of them which make child health care a relatively attractive and promising candidate for evaluation.

There are aspects of methodologic boundaries, however, which are more troublesome than underutilization of good methods. I am not competent to discuss the intricacies and specifics of many of these issues, but the outlines of the debate seem to be as follows. There are those who argue that rigorous experimental research methodologies, in particular random assignment of control and experimental groups, not only can but should be utilized whenever possible (a circumstance much more common than detractors of the methods claim), because they provide the most reliable and valuable data for policy making.

The major function of a controlled randomized experiment is to obtain data in a way which is conducive to making a judgment about the efficacy of a treatment program. Experimental appraisal of social programs can therefore be reasonably expected yield information which is useful for decisions in

a more direct and unequivocal way than other data collection strategies. (Boruch, 1972, p. 1)

Campbell suggests an ethical, as well as a scientific rationale for such rigorous designs. He calls for a wider

recognition of randomization as the most democratic and moral means of allocating scarce resources (and scarce hazardous duties), plus the moral imperative to further utilize the randomization so that society may indeed learn true value [sic] of the supposed boon. This is the ideology that makes possible "true experiments" in a large class of social reforms. (Campbell, 1969; in Caro, 1971, p. 248)

We need to know more about the limitations of experimental and quasi-experimental designs in specific situations (c.f. Campbell & Stanley, 1963), and about the possibilities of achieving randomized treatment and control groups. In particular, we need to know whether experimental techniques can work for program evaluation which must have policy implications, for Boruch's optimism is not widely shared. In a methodologic review of the Coleman Report on compensatory education programs, Cain and Watts make a two-pronged critique:

Little or no theoretical justification is offered for the election of explanatory variables, for their functional form, or for the inclusion or exclusion of variables under different specifications of the model...We have no way of knowing, for example, whether a variable directly represents a policy instrument or is only indirectly related to policy through some other unmeasured (or partially measured) relationships.

...Second, the criterion used...to assess or evaluate the statistical performance of the variables is inappropriate. Instead of providing information about the quantitative effect of a variable in altering educational achievement --information which would enable the reader to assess the feasibility and costliness of operating on the variable-- the Report provides information about a statistical measure of the variable's performance...which gives no clear guidance for translating the statistical findings into policy action. (Cain & Watts, 1970, p. 229)

As I have discussed, methodologies can always be attacked by critics of findings or of programs. What needs to be done more carefully

(Boruch's work is a hopeful step in the right direction) is a sorting out of real issues and weaknesses from those which may be academically real but which do not affect the policy implications of evaluations.

Leaving aside methodological limitations, one realizes in studying the definitions and descriptions of evaluation the degree to which both theory and practice have remained aloof from and leery of real decision-making situations. Part of the problem is an insufficient conceptual base.

One of the major reasons evaluation is in difficulty is that knowledge of the decision-making process and of the methodologies for relating evaluation to decision-making is woefully inadequate. (Stufflebeam, et al., 1971, p. 16)

This is no trivial inadequacy when one realizes that evaluation, as opposed to social research, was developed as a key input into decision-making.

Evaluation is a decision-making tool. Its success or failure must be measured therefore in terms of its impact on changing program policies and resources allocations. (Wholey, et al., 1970, p. 46)

Concern with policy implications highlights the second, or ideologic limit on evaluation: its proper role in assessing the overall importance and ethical acceptability of programs. Here the common stress on the value-free nature and scientific objectivity of evaluation (and indeed policy analysis techniques in general) tends to blur crucial questions, to mask problems in apparent success by partial statistical measures, or to create failures where none exist. This latter situation has been discussed at length in the text. We need to be more aware of the interactions of evaluation with other political system variables. Advocacy and actionability evaluations are an attempt to do this. The

other side of the coin, the over-generalization from incomplete measures has not been as thoroughly discussed. Once "progress" can be shown, it becomes more difficult for reform movements to keep pressure on administrators or legislators for further action. We Americans have a great faith in our ability to solve problems and a short attention span for complex, half-finished business. The Citizens' Board of Inquiry into Hunger captures the essence of the problem in their assessment of the progress made toward relieving hunger in America in the four years since their first report:

If we were reviewing here a matter such as increased highway construction, or any other materially-centered government project in which statistics provide the only measure of success, failure or commitment, we would take pride in this data, and look with confidence and composure at the job still to be done.... But we are considering something infinitely more complex, more profound. We are considering hunger and its debilitating effects on human personality, hunger and development, considerations deserving the highest priority in a civilized nation. And, what is at issue, as much as the will of this administration, or any administration, to take action, is the humanity of and the swift ability of our methods. (Citizens' Board of Inquiry, 1972, p. 9)

The very bases of evaluation for complex social programs such as those we are concerned with may be unique. If this is so, then transferring methods either from social science research (where only single variables or controlled sets of a few variables are generally studied) or from management and operations research (where non-human resources and non-social goals are manipulated) may be both an inappropriate and indeed a dangerous overstepping of boundaries. Campbell suggests that the political dangers of evaluation ("Specific reforms are advocated as though they were certain to be successful. For this reason, knowing outcomes has immediate political implications.")

and also its tendency to kill off reform with negative findings could be eased by a shift in political posture

from the advocacy of a specific reform to the advocacy of the seriousness of the problem, and hence to the advocacy of persistence in alternative reform efforts should the first one fail. (Campbell, 1969, p. 234)

We should be especially concerned with methodologies or models which separate ethical and pragmatic or scientific measures: "Humane grounds (social value) for providing care are not invalid grounds. They merely fall outside the purview of the special competence of the economist." (Herbert Klarman; quoted in Sparer & Johnson, 1971, p. 932) If the economist (or the psychologist or the sociologist or the planner) is not going to claim competence to elucidate the social value of the social programs he studies, then they will have to be directed by someone who can interject that perspective. But policy analysis is replacing raw politics as the socially acceptable ideology of power, and policy science (be it PPBS or another variant) also pretends to be value-free, to make objective, rational choices among alternatives on clear grounds of evidence or cost and benefit. The individual moral judgment of the voter is rarely able to express itself in the enormity and administrative anonymity of modern government. Are we truly ready to transcend morality and ethics? Or are we only neglecting them?

To discuss evaluation any further, the boundaries of knowledge itself must be examined. For evaluation is inextricably tied to our cultural beliefs in the power and indeed the necessity of scientific knowledge. We hate above all else not to know, not to be able to make the world explicable, replicable, and preferably graphable as well.

That which we cannot measure or reproduce we put out of mind. In medical science "miasma theory", vitamin E, faith-healing, and many psychosomatic complaints are considered less than real; only those diseases, syndromes, treatments, and cures approved by the A.M.A. and amenable to research at NIH are granted full public acknowledgment.* The scientific approach becomes a "dichotomization of life and knowledge, knowing and living," based on the notion that

there is an objective world with knowable laws, that can be determined through various empirical testing procedures. Given knowledge of these laws and facts, man can then decide how to live. They are assumed to be separable: scientific knowledge first, then morality....Science implies that man learns about his environment only through objective study.
(Ted Butz; quoted in Sackrey, 1973, p. 75)

Wolin discusses the correlates of this split in political theories of organization and bureaucracy, or what he terms "the sublimation of politics." Referring to one of the fathers of modern organizational theory he captures the dreams of scientific rationalism in public affairs:

Saint-Simon's vocabulary "organization" connoted far more than a simple condition of social harmony and political stability. Organization promised the creation of a new structure of power...."Men shall henceforth do consciously, and with better directed and more useful effort, what they have hitherto done unconsciously, slowly, indecisively, and too ineffectively. (Wolin, 1960, p. 377)

In a rather frightening essay entitled "The Cybernetic State," Allen Schick (1970) explores the implications of the post-bureaucratic stage of government, when systematic rather than functional form is

* Adair and Deuschle (1970) provide a fascinating account of the more reasonable ways of the Navajo of accepting and utilizing modern medicine. In cases where it works, such as drug treatment for tuberculosis, the medicine men have acknowledged the superiority of modern medicine and send people to hospitals; where it has no greater success than traditional medicine, as for psychiatric disorders or chronic conditions, they have held their ground--and their patients.

adopted for the goal-directed government, which acts as a "servomechanism, concerting the polity and the economy to achieve public objectives." In this world, planners and systems engineers have replaced those with functional specialties (whose "interests get in the way of the results") and the right of free speech has become gradually eroded as the speaker has "little ability to challenge the dominance of the experts and... little ability to sway the course of public policy [much of which becomes automatically linked to changes in social indicators] by his vote." (p. 278) He notes the increased overlap between public and private spheres.

For example, can the state compel an individual not to smoke? Can it require adults to work? Can it force mothers to send three-year-olds to school? None of these issues is distinctly new; but what is new is the ground on which they are being fought and will be decided. Prohibition was a moral issue; but smoking is already a scientific one. (Schick, 1970; in Gelb & Palley, 1971, p. 279)

With aminocentesis and related techniques already enabling doctors to detect many fetal abnormalities, the issues are not in the future but in the present. Can the state compel the pregnant woman to have an abortion if the family will be unable to support a multiply-handicapped or retarded child, or if chromosomal patterns associated with criminality are found? What is frightening about the reality of such decisions (for the woman if not for the state) is that they are so often couched in scientific jargon which hides the moral complexities involved. I don't have much in common with opponents of abortion reform, but I deeply respect some of them because they at least recognize that the argument is a moral and ethical one, not one which is purely political, social, or reducible to hard evidence on the safety of the

procedures or the numbers of illegitimate or unwanted births each year.

To practice "normal science" is to be tempted into a personal style whereby one increasingly narrows one's range of inquiry and increasingly refuses to consider the moral implications of what one is doing. (Sackrey, 1973, p. 76)

In such cases, too much knowledge can be a dangerous thing.

But there are few simple, untrained minds among us, and they have little voice. Many of us have been trained too well, it seems. We lead our lives as if we only knew what the tests show. Our minds are anchored to statistics, and so to the present....

We love our technical solutions so deeply that we refuse to acknowledge the over-riding problems they create. We know what we know with such allegiance that our very loyalty betrays our purpose. (Gerzon, 1973)

When we go this far in our reliance upon and faith in our powers of knowledge and certainty, then the flowers have become weeds. Science should not blind us to our own humanity, to the trustworthiness of our own eyes and ears, to our unaugmented powers of reason, or to the irreducible complexities and uncertainties of life. While not all of us possess the opportunity or the empathetic perspicacity of a Jack London or a Bobby Kennedy, we need to develop those latent talents in ourselves before they indeed atrophy as we know the unused intellect does.

The present work has certainly not provided answers as to how to "systematically" go about doing this. But part of the process must certainly be a more thoughtful analysis and evaluation of not only programs of social action but also the means by which those programs come to be, come to die, and by which we learn about them. That analysis will not be simple or conclusive. It is paradoxical to expect to know precisely the limits of our knowledge. We need to be more tolerant of and sensitive to the boundaries of social institutions and concepts, without

falling into the existential despair of saying that nothing matters or is any better than anything else. Child development, for example, is a powerful synthesizing and intellectually-expanding construct for understanding and arguing for children and their needs. But it can also be a limiting idea, if by depending on it too heavily, we slight pressing needs for which we can demonstrate no future developmental impact. Likewise, evaluation techniques are an essential part of modern society, but they can be made trivial, ludicrous, or life-threatening by misuse. If evaluation is used in the service of reason and not as its technological or cybernetic replacement, it is a powerful means of acquiring knowledge, or monitoring the goodness of our social assumptions and actions, and of revealing the value conflicts (or consensus) which lie behind them.

I am not trying to argue for "moderation in all things," but rather for a sensitivity and depth of understanding which obviates the need for such quickie rules-of-thumb: for a perpetual openness of the mind to the full range of possibilities and implications of human social endeavors. Nick Kotz (1969) ends his absorbing and shocking expose of the politics of hunger in America with some words which have a balanced (but not middle-of-the-road) quality rare in muck-raking journalism and which express the type of approach I am seeking.

Perhaps these questions have to be asked because we have accepted too much at face value about ourselves; we have accepted too many assumptions about the essential benevolence of our institutions, without really analyzing how they really operate and how they affect all Americans.

The politics of hunger in America is a dismal story of human greed and callousness, of immorality sanctioned and aided by the government of the United States. But it

is also a story that does provide hope that men can change things, that men do care about fulfilling this country's highest ideals and do care about their fellow human beings. (Kotz, 1969, p. 246)

To the extent that this openness and honesty of analysis take place, planning and humanitarian goals need not be antithetical and indeed may enrich and foster each other.

I am told that as a young child I loved to help tend my family's vegetable and flower gardens. My technique was to pull up any questionable plants (they being numerous as my horticultural knowledge was limited) and bring them to my mother or father, asking, "Dis a weed?" At which point the question was moot; whatever it was was dead. I have become more selective in my uprooting in more recent years and hence wreck less havoc on the tulip population. But I still think that the question is a damn good one to ask.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Adair, J., & Deuschle, K.W. The People's Health: Medicine and anthropology in a Navajo community. New York: Appleton-Century-Crofts, 1970.
- Alpert, J.J., Heagarty, M.C., Robertson, L., Kosa, J., & Haggerty, R.J. Effective use of comprehensive pediatric care: Utilization of health resources. American Journal of Diseases of Children 116: 529-33 (1968).
- Alpert, J.J., Kosa, J., Haggerty, R.J., Robertson, L.S., & Heagarty, M.C. Attitudes and satisfactions of low-income families receiving pediatric care. American Journal of Public Health 60: 499-506 (1970).
- American Academy of Pediatrics. Lengthening Shadows: A report of the Council on Pediatric Practice of the Academy on the delivery of health care to children, 1970. Evanston, Illinois: The Academy, 1971.
- American Public Health Association. Glossary of administrative terms in public health. American Journal of Public Health 50: 225-226 (1960).
- Apgar, V. A proposal for a new method of evaluation of the newborn infant. Anesthesia and Analgesia 32: 260 (1953).
- Appalachian Regional Commission. Programs for Infants and Young Children: Part III. Health. Washington, D.C.: The Commission, 1970.
- Arnold, F.A., Dean, H.T., Jay, P., & Knutson, J.W. Effect of fluoridated public water supplies on dental caries prevalence. Public Health Reports 71: 652-58 (1956).
- Ast, D.B. Effectiveness of Water fluoridation. Journal of the American Dental Association 65: 581 (1962).
- Bailey, R.M. Philosophy, faith, fact and fiction in the production of medical services. Inquiry, Vol. VII, No. 1 (March 1970), pp. 37-53.
- Bellin, S.S., & Geiger, H.J. Actual public acceptance of the neighborhood health center by the urban poor. Journal of the American Medical Association 205: 663-669 (1968).
- Beloff, J.S., & Weinerman, E.R. Yale studies in family health care: I. Planning and pilot test of a new program. Journal of the American Medical Association 199: 383-389 (1967).
- Beloff, J.S., & Willet, M. Yale studies in family health care: III. The family care team. Journal of the American Medical Association 205: 663-669 (1968).

- Bergman, T. Children in the Hospital. New York: International University Press, 1965.
- Birch, H.G. Health and education of socially disadvantaged children. Developmental Medicine and Child Neurology 10: 580-99 (1968).
- Birch, H.G., & Gussow, J.D. Disadvantaged Children: Health, nutrition, and school failure. New York: Harcourt, Brace, and World, 1970.
- Bloom, B.L. The "medical model," miasma theory, and community mental health. Community Mental Health Journal 1: 333-338 (1965).
- Boruch, R.F. Problems in research utilization: Use of social experiments; of experimental results; and of auxiliary data in experiments. Annals of the New York Academy of Sciences. Winter, 1972 (forthcoming).
- Breslow, L. Some essentials in a national program for child health. Pediatrics 44: 327-332 (1969).
- Cain, G.C., & Watts, H.W. Problems in making policy inferences from the Coleman report. American Sociological Review 35: 228-242 (1970).
- Callahan, D.L. Focus on nutrition--You can't teach a hungry child: Part II. School Food Journal, September, 1971, p. 25.
- Campbell, D.T. Reforms as experiments. American Psychologist 24: 409-429 (1969).
- Campbell, D.T., & Erlebacher, Albert. How regression artifacts in quasi-experimental evaluations can mistakenly make compensatory education look harmful. In Hellmuth, J. (Ed.) Disadvantaged Child, Vol 3. Compensatory Education: A national debate. New York: Brunner/ Mazel, 1970.
- Campbell, D.T., & Stanley, J.C. Experimental and Quasi-experimental Designs for Research. Chicago: Rand-McNally & Company (for the American Educational Research Association), 1963.
- Caro, F.G. (Ed.) Readings in Evaluation Research. New York: Russell Sage Foundation, 1971.
- Chase, H.P., & Martin, H.P. Undernutrition and child development. New England Journal of Medicine 282: 933-38 (1970).
- Child Health Care Policy Task Force. Summary Report of the Task Force. Unpublished paper prepared for the Office of Child Development, HEW. April 7, 1972.
- Children and Youth projects: A progress report. Maternal and Child Health Information, No. 14, June, 1971, pp. 1-7.
- Christakis, G., et al. A nutritional epidemiologic investigation of 642 New York City children. American Journal of Clinical Nutrition 21: 107-26 (1968).

Citizens' Board of Inquiry into Hunger and Malnutrition in the United States. Hunger U.S.A. Revisited: A report by the Board. Published in cooperation with the National Council on Hunger and Malnutrition and the Southern Regional Council, 1972.

_____. Hunger U.S.A.: A report by the Board. Washington, D.C., The New Community Press, 1968.

Close, K. Promoting child health through comprehensive care. Children 16: 130-137 (1969).

Collver, A., Have, T.R., & Speare, M. Factors influencing the use of maternal health services. Social Science and Medicine 1: 293-305 (1967).

Colombo, T.J., Saward, E.W., & Greenlick, M.R. The integration of an OEO health program into a prepaid comprehensive group practice plan. American Journal of Public Health 59: 641-50 (1969).

Committee on Maternal Nutrition. Food and Nutrition Board. (Ed.) Maternal Nutrition and the Course of Pregnancy. Washington, D.C.: National Research Council, National Academy of Sciences, 1970.

de Geyndt, W. Five approaches for assessing the quality of care. Hospital Administration, Winter, 1970, pp. 21-42.

de Geyndt, W. A tentative framework for the evaluation of the Children and Youth Program. Paper prepared for Children and Youth Evaluation Conference, June 17, 1969. Systems Development Project Comment Series, No. 9-7 (20), 1969.

de Vise, P. Persistence of Chicago's dual hospital system. In de Vise, P., et al. Slum Medicine: Chicago's apartheid health system. Chicago: University of Chicago, Community and Family Study Center, 1969.

Dimock, G.H. The Child in Hospital: A study of his emotional and social well being. Philadelphia: F.A. Davis, 1960.

Donabedian, A. An evaluation of prepaid group practice. Inquiry, Vol. VI, No. 3, (1969), pp. 3-27.

_____. Promoting quality through evaluating the process of patient care. Medical Care 6: 181-202 (1968).

Dorfman, Robert (ed.) Measuring Benefits of Government Investments. Washington, D.C.: The Brookings Institution, 1965.

Drew, E.B. The health syndicate: Washington's noble conspiracy. The Atlantic, Vol 220, No. 6 (December, 1967), pp. 75-82.

Drillien, C.M. The Growth and Development of the Prematurely Born Infant. Baltimore: Williams and Wilkins, 1964.

- Drillien, C.M. The incidence of mental and physical handicaps in school-age children of very low birthweight. Pediatrics 27: 452-464 (1961).
- Dubos, R. The Mirage of Health: Utopias, progress and biological change. New York: Harper & Row, 1959.
- Duff, R.S., & Hollingshead, A.B. Sickness and Society. New York: Harper & Row, 1968.
- Duhl, Leonard J., & Leopold, R.L. Mental Health and Urban Social Policy. San Francisco: Jossey-Bass, Inc., 1969.
- Edgerton, J.W. Evaluation in community mental health. In American Psychological Association. Task Force on Community Mental Health, Division 27. Issues in Community Psychiatry and Preventive Mental Health. New York: Behavioral Publications, Inc., 1971. Pp. 89-107.
- Ehrenreich, B., & Ehrenreich, J. American Health Empire: Power, profits, and politics. A Health-PAC book. New York: Random House, 1970.
- Elias, M. Behavioral consequences of malnutrition in infancy: A review of human studies. Unpublished paper, Department of Nutrition, School of Public Health, Harvard University, 1971.
- Elinson, J., & Herr, C.E.A. A sociomedical view of neighborhood health centers. Medical Care 8: 97-103 (1970).
- Etzioni, A. Two approaches to organizational analysis: A critique and a suggestion. Administrative Science Quarterly 5: 257-278 (1960). Reprinted in Schulberg, et al., op. cit. (1969), pp. 101-120.
- Fanshel, S., & Bush, J.W. A health-status index and its application to health-services outcomes. Operations Research 18: 1021-66 (1970).
- Fomon, S. Prevention of iron-deficiency anemia in infants and children of preschool age. Public Health Service Publication No. 2085. Washington, D.C.: Government Printing Office, 1970.
- Food Research and Action Center. If We Had Ham, We Could Have Ham and Eggs, If We Had Eggs: A study of the National School Breakfast Program. New York: Gazette Press, 1972.
- Freeman, H.E., & Sherwood, C.C. Research in large-scale intervention programs. Journal of Social Issues 21 (1): 11-28 (1965). Reprinted in Schulberg, et al., op. cit., pp. 73-91.
- Freidson, E. Patients' Views of Medical Practice. New York: Russell Sage Foundation, 1961.
- Garfield, S.R. The delivery of health care. Scientific American 222: 15-23 (1970).

- Garvue, Flanagan, & Castine. The National School Food Service and Nutrition Education Finance Project. Special Study No. 8, National Education Finance Project. Washington, D.C.: U.S. Department of Agriculture, 1971.
- Geomet, Inc. Study to evaluate the BEO Neighborhood Health Center program at selected centers. Vols. I - III. Draft of final report prepared for Office of Economic Opportunity. Rockville, Md.: Geomet, Inc., October 1, 1971.
- Gerzon, M. A Childhood for Every Child: Toward a politics of parenthood. New York: Outerbridge & Lazard, 1973. (In press.)
- Gil, D.G. Violence against children: Physical abuse in the United States. Cambridge, Massachusetts: Harvard University Press, 1970.
- Glass, L., Kolko, N., & Evans, H. Factors influencing predisposition to serious illness in low birth weight infants. Pediatrics 48: 368-71 (1971).
- Goodrich, C.W., Olendzki, M.C., & Reader, G.G. Welfare Medical Care: An Experiment. Cambridge, Massachusetts: Harvard University Press, 1970.
- Gordis, L., & Markowitz, M. Evaluation of the effectiveness of comprehensive and continuous pediatric care. Pediatrics 48: 766-776 (1971).
- Greenberg, I.G., & Rodburg, M.L. The role of prepaid group practice in relieving the medical care crisis. Harvard Law Review 84: 887-1001 (1971).
- Gussow, J.D. Bodies, brains, and poverty: Poor children and the schools. IRCD Bulletin, Vol. VI, No. 3 (September, 1970), pp. 3-4+.
- Gutelius, M.F. The problem of iron deficiency anemia in preschool Negro children. American Journal of Public Health 59: 290-95 (1969).
- Haggerty, R.J. The boundaries of health care. Pharos 35: 106-111 (1972).
- _____. Present strengths and weaknesses in current systems of comprehensive health services for children and youth. American Journal of Public Health 60: 74-98 (1970) (A).
- _____. Research issues in child health: A Head Start research seminar: II. Some medical and economic issues. Pediatrics 45: 702-12 (1970) (B).
- Harper, D., & Babigan, H. Evaluation research: The consequences of program evaluation. Mental Hygiene 55: 151-156 (1971).
- Hellmuth, J. (Ed.) Disadvantaged Child. 3 vols. New York: Brunner, Mazel, 1967.

- Huron Institute. Federal Programs for Young Children; Review and Recommendations. 3 vols. Final report to HEW, September 15, 1972.
- Hunt, D. Health services to the poor: A survey of the prevalence and causes of infant mortality in the U.S. Inquiry 6: 27-38 (Dec., 1969).
- Hunt, E. Infant mortality trends and maternal and infant care. Children 17: 88-90 (1970).
- Hunter, G. Health care through Head Start. Children 17: 149-153 (1970).
- Iglehart, J.K. Budget report/ HEW Department, largest federal spender, seeks to funnel more money to the poor. National Journal, January 29, 1972, pp. 168-188.
- Illsley, R. Early prediction of perinatal risk. Proceedings of the Royal Society of Medicine 59: 181-184 (1966).
- International Encyclopedia of the Social Sciences. 17 vols. David L. Sills, ed. New York: Macmillan Company and The Free Press, 1968.
- James, G. Evaluation in public health practice. American Journal of Public Health 52: 1145-1154 (1962). Reprinted in Schulberg, et al., op. cit., pp. 29-41.
- Joint Commission on the Mental Health of Children. Crisis in Child Mental Health: Challenge for the 1970's. New York: Harper & Row, 1970.
- Kaplan, B.J. Malnutrition and mental deficiency. Psychological Bulletin 78: 321-334 (1972).
- Kempe, C.H., Silver, F., Steele, B.F., Droegemueller, W., & Silver, J.K. The battered-child syndrome. Journal of the American Medical Association 181: 17-24 (1962).
- Kerr, G.R. The nutritional correlates of life: Growth and learning. Journal of School Health 42: 191-96 (1972).
- Kerr, M., & Trantow, D.J. Defining, measuring, and assessing the quality of health services. Public Health Reports 84: 415-24 (1969).
- Knobloch, H., & Pasamanick, B. Prospective studies on the epidemiology of reproductive casualty: Methods, findings, and some implications. Merrill-Palmer Quarterly of Behavior and Development 12: 27-43 (1966).
- Kosa, J., Antonovsky, A., & Zola, I.K. (Eds.) Poverty and Health: A sociological analysis. Cambridge, Massachusetts: Harvard University Press, 1969.
- Kotz, N. Let Them Eat Promises: The politics of hunger in America. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1969.

- Lesser, A.J. The federal government in child health care. Pediatric Clinics of North America 16: 891-900 (1969).
- _____. Closing gaps in the nation's health services for mothers and children. Bulletin of the New York Academy of Medicine 40: 1249-1253 (1965).
- Levin, A.L. Cost-effectiveness in maternal and child health: Implications for program planning and evaluation. New England Journal of Medicine 278: 1041-47 (1968).
- Lowe, C.U. The role of government in implementing good nutritional practice. Bulletin of the New York Academy of Medicine 47: 647-56 (1971).
- Lyden, F.J., & Miller, E.G. (Eds.) Planning-Programming-Budgeting: A Systems Approach to Management. 2nd. edition. Chicago: Markham Publishing Company, 1972.
- MacMahon, B., Pugh, T., & Hutchison, G. Principles in the evaluation of community mental health programs. American Journal of Public Health 51: 963-68 (1961). Reprinted in Schulberg, et al., op. cit., pp. 51-58.
- Madison, D.L. Organized health care and the poor. Medical Care Review 26: 783-807 (1969).
- Marris, P., & Rein, M. Dilemmas of Social Reform: Poverty and community action in the United States. New York: Atherton Press, 1969.
- McDermott, W., Deuschle, K.W., & Barnett, C.R. Health care experiment at Many Farms. Science 175: 23-31 (1972).
- McKeown, T. Medicine in Modern Society: Medical planning based on evaluation of medical achievement. London: G. Allen and Unwin, 1965.
- Menken, J. The health and social consequences of teenage childbearing. Family Planning Perspectives 4: 45-53 (1972).
- Minnesota Research Systems, Inc. Summary Analysis of Present and Future Met and Unmet Health Needs of Children. Report prepared for HEW Child Health Care Policy Task Force, February, 1972.
- Monahan, H.B., & Spencer, E.C. Deterrents to prenatal care. Children 9: 114-119 (1962).
- Morehead, M.A., Donaldson, R.S., & Seravalli, M.R. Comparisons between OEO neighborhood health centers and other health care providers on ratings of the quality of health care. American Journal of Public Health 61: 1294-1306 (1971).
- Moriyama, I.M. Problems in the measurement of health status. In Sheldon, E.B., & Moore, W.E. (Eds.) Indicators of Social Change: Concepts and measurements. New York: Russell Sage Foundation, 1968.

- National Academy of Sciences. Health Services Research Study, Institute of Medicine. Development of Methodology for Evaluation of Neighborhood Health Centers. Prepared for Health Evaluation Office, HEW. Washington, D.C.: The Academy, February, 1972.
- National Advisory Commission on Health Manpower. Report of the Commission. 2 Vols. Washington, D.C.: Government Printing Office, 1967.
- National Institute of Mental Health. Evaluation in Mental Health. Public Health Services Publication No. 413. Washington, D.C.: Government Printing Office, 1955.
- Nay, J.N., Vogt, L.M., & Wholey, J.S. Health Start: Interim analysis and report. Working Paper No. 961-2. Washington, D.C.: The Urban Institute, January 3, 1972.
- Newberger, Eli. Study of social illness in children: Hospital-based program for research and intervention into the determinants of distressed parent-child relationships and their developmental consequences. Unpublished grant application. November, 1971.
- North, A.F. Research issues in child health: A Head-Start research seminar. I. Introduction and overview. Pediatrics 45: 690-701 (1970).
- _____. Project Head Start and the pediatrician. Clinical Pediatrics 6: 191-194 (1967).
- Notkin, H., & Notkin, M.S. Community participation in health services: A review article. Medical Care Review 27: 1178-1201 (1970).
- Owen, G.M., & Kram, K.M. Nutritional status of preschool children in Mississippi: Food sources of nutrients in the diets. Journal of the American Dietetic Association 54: 490-94 (1969).
- Paige, D.M. School feeding program: Who should receive what? Journal of School Health 41: 261-3 (1971) (A).
- _____. The school lunch program: An underachiever. Unpublished paper presented at Convention of American School Health Association, October, 1971. (B).
- Pasamanick, B., & Knobloch, H. Retrospective studies of the epidemiology of reproductive casualty: Old and new. Merrill-Palmer Quarterly of Behavior and Development 12: 17-26 (1966).
- _____. Brain damage and reproductive casualty. American Journal of Orthopsychiatry 30: 298-305 (1960).
- Patterson, P.K., et al. Parent reaction to the concept of pediatric assistants. Pediatrics 44: 69-96 (1969).

- Peterson, M.L. The first year in Columbia: Assessment of low hospitalization rate and high office use. Johns Hopkins Journal of Medicine 128: 15-23 (1971).
- Perrott, G.S., & Chase, J.C. The Federal Employees Health Benefits Program. Group Health and Welfare News, Special Supplement, 1968.
- Read, M.S. The biological bases: Malnutrition and behavioral development. In National Society for the Study of Education. 1972 Yearbook on Early Childhood Education. In press.
- _____. Malnutrition and mental development: Needed research to clarify critical questions. Special report prepared for National Institute of Child Health and Human Development. Bethesda, Md.: National Institutes of Health, May, 1969.
- Richmond, J.B., & Weinberger, H.L. Program implications of new knowledge regarding the physical, intellectual and emotional growth and development and the unmet needs of children and youth. American Journal of Public Health 60: Supplement (April, 1970), 23-73.
- Rider, R.B., Tayback, M., & Knobloch, H. Association between premature birth and socio-economic status. American Journal of Public Health 45: 1022-1028 (1955).
- Rivlin, A. Systematic Thinking for Social Action. Washington, D.C.: The Brookings Institution, 1970.
- Robinson, N.M., & Robinson, H.B. A follow-up study of children of low birthweight and control children at school age. Pediatrics 35: 425-433 (1965).
- Rossi, P.H., & Williams, W. (Eds.) Evaluating Social Programs: Theory, Practice, and Politics. New York: Seminar Press, 1972.
- Rowe, D.S., Leonard, M.F., et al. A hospital program for the detection and registration of abused and neglected children. New England Journal of Medicine 282: 950-52 (1970).
- Rutter, M., Tizard, J., & Whitmore, K. (Eds.) Education, Health and Behavior: Psychological and medical study of child development. New York: John Wiley & Sons, Inc., 1970.
- Sackrey, C. The Political Economy of Urban Poverty. New York: W.W. Norton & Company, Inc., 1973.
- Sandstead, H.H., Carter, J.P., House, F.R., et al. Nutritional deficiencies in disadvantaged preschool children: Their relationship to mental development. American Journal of Diseases of Children 121: 455-63 (1971).

- Sartwell, P.E. (ed.) Maxcy-Rosenau Preventive Medicine and Public Health. Ninth Edition. New York: Appleton-Century-Crofts, 1965.
- Saward, E.W., Blank, J.D., & Greenlick, M.R. Documentation of twenty years of operation and growth of a prepaid group practice plan. Medical Care 6: 251-44(1968).
- Schachter, F.F., & Apgar, V. Perinatal asphyxia and psychologic signs of brain damage in childhood. Pediatrics 24: 1016-1025 (1959).
- Schick, A. The cybernetic state. Trans-Action, February, 1970, pp. 15-26. Reprinted in Gelb, J., & Palley, M.F. (Eds.) The Politics of Social Change: A reader for the seventies. New York: Holt, Rinehart and Winston, Inc., 1971; pp. 269-282.
- Schulberg, H.C., Sheldon, A., & Baker, F. (Eds.) Program Evaluation in the Health Fields. New York: Behavioral Publications, 1969.
- Scrimshaw, N.S. Evaluation of nutrition in pregnancy. Journal of the American Dietetic Association 26: 21-24 (1950).
- _____. Malnutrition and infection. Borden Review of Nutritional Research 26: 17-29 (1965).
- Segal, J.A. Food for the Hungry: The reluctant society. Baltimore: The Johns Hopkins University Press, 1970.
- Shaeffer, H.R., & Callender, W.M. Psychological effects of hospitalization in infancy. Pediatrics 24: 528- 539 (1959).
- Shapiro, S. End result measurements of quality of medical care. Milbank Memorial Fund Quarterly, Vol. XLV, No. 2, Part 1 (April 1967), pp. 7-30.
- Shapiro, S., Jacobziner, H., Densen, P.M., & Weiner, L. Further observations on prematurity and perinatal mortality in a general population and in the population of a prepaid group practice medical care plan. American Journal of Public Health 50: 1304-1317 (1960).
- Shapiro, S., Schlesinger, E.R., & Nesbitt, R.E.L. Infant, perinatal, maternal and childhood mortality in the United States. Cambridge, Massachusetts: Harvard University Press, 1968.
- Shore, M.F. (Ed.) "Red is the color of hurting": Planning for children in the hospital. Bethesda, Md.: National Institute of Mental Health, 1967.
- Siegel, E. Migrant families: Health problems of children. Clinical Pediatrics 5: 635-640 (1966).
- Silver, G.A. Family Medical Care: A report on the Family Health Maintenance Demonstration. Cambridge, Mass.: Harvard University Press, 1963.

- Silver, H.K., & Kempe, C. The problem of parental criminal neglect and severe physical abuse of children. American Journal of Diseases of Children 98: 528 (1959).
- Silver, H.K., Ford, L.C., & Day, L.R. The pediatric nurse practitioner program: Expanding the role of the nurse to provide increased health care for children. Journal of the American Medical Association 204: 298-302 (1968).
- Simons, B., Downs, E.F., Hurster, M.N., & Archer, M. Child abuse--Epidemiologic study of medically reported cases. New York Journal of Medicine 66: 2783-2788 (1966).
- Smith, H.D. Pediatric lead poisoning. Archives of Environmental Health 8: 256-261 (1964).
- Somers, A. Health Care in Transition. Chicago: Hospital Research and Educational Trust, 1971.
- Sparer, G., & Johnson, J. Evaluation of OEO neighborhood health centers. American Journal of Public Health 61: 931-942 (1971).
- Spitz, R.A. Hospitalism: An inquiry into the genesis of psychotic conditions in early childhood. In Fenichel, O., et al. (Eds.) Psychoanalytic Study of the Child, Vol. 1. New York: International Universities Press, 1945.
- _____. Hospitalism: A follow-up report. Psychoanalytic Study of the Child, Vol. 2, (1946), pp. 113-117.
- Stansky, P., & Abrahams, W. The Unknown Orwell. New York: Alfred A. Knopf, 1972.
- Stein, Zena, Susser, M., et al. Nutrition and mental performance: Prenatal exposure to the Dutch famine of 1944-45 seems not related to mental performance at age 19. Science 178: 708-713 (1972).
- Stoeckle, J.D., & Candib, L.M. The neighborhood health center--reform ideas of yesterday and today. New England Journal of Medicine 280: 1385-91 (1969).
- Strauss, A.L. Medical ghettos. Trans-action, Vol. 4, No. 6 (1967), pp. 7-15.
- Strauss, M.A., & Sparer, G. Basic utilization experience of OEO comprehensive health services projects. Paper presented to the Health Planning Processes and Techniques Section, American Public Health Association, Houston, October, 1970. Reprinted by the Office of Health Affairs, Office of Economic Opportunity.
- Stufflebeam, D.I., et al. Educational Evaluation and Decision-Making. Itasca, Illinois: F.E. Peacock, 1971.

- Suchman, E.A. Evaluative Research: Principles and practice in public service and social action programs. New York: Russell Sage Foundation, 1967.
- Sullivan, D.F. Conceptual Problems in Developing an Index of Health. Vital and Health Statistics, National Center for Health Statistics, Series 2, No. 17. Public Health Service Publication No. 1000. Washington, D.C.: Government Printing Office, May, 1966.
- Talbot, N.B. Has psychologic malnutrition taken the place of rickets and scurvy in contemporary pediatric practice? Pediatrics 31: 909-918 (1963).
- Talbot, N.B., & Howell, M. Social and behavioral causes and consequences of disease among children. In Talbot, N.B., Kagen, J., & Eisenberg, L. (Eds.) Behavioral Science in Pediatric Medicine. Philadelphia: W.B. Saunders, 1971.
- Thorner, R.M. Health program evaluation in relation to health programming. HSMHA Reports 86: 525-532 (1971).
- Tompkins, W.T., Mitchell, R.M., & Wiehl, D.G. Maternal and newborn nutrition studies at Philadelphia Lying-In Hospital: Prematurity and maternal nutrition. In The Promotion of Maternal and Newborn Health. New York: Milbank Memorial Fund, 1955.
- United States Congress. 91st Congress, First Session. Joint Economic Committee. Subcommittee on Economy in Government. The Analysis and Evaluation of Public Expenditures: The PPB System. A Compendium of papers submitted to the Subcommittee. (The Proxmire Subcommittee) Washington, D.C.: Government Printing Office, 1969.
- United States Department of Agriculture. Dietary Levels of Households in the United States, Spring, 1965: A preliminary report. Agriculture Research Service Publication. Washington, D.C.: Government Printing Office, 1968.
- United States Department of Health, Education, and Welfare. Maternal and Child Health Service. Promoting the Health of Mothers and Infants. Washington, D.C.: Dept. of HEW, 1970.
- _____. Toward a Social Report. Washington, D.C.: Government Printing Office, 1969.
- _____. Perspectives on Human Deprivation: Biological, psychological, and sociological. HEW, Public Health Service, National Institute of Child Health and Human Development, 1968.
- _____. Office of the Assistant Secretary for Program Coordination. Program Analysis: Maternal and Child Health Care Programs. Washington, D.C.: HEW, October, 1966.

- United States Office of Management and Budget. Special Analyses, Budget of the United States Fiscal Year 1973. Washington, D.C.: Government Printing Office, 1972.
- Vernon, M. Prematurity and deafness: The magnitude and nature of the problem among deaf children. Exceptional Children 30: 289-298 (1967).
- Wagner, M.G. Research issues in child health: A Head Start research seminar: III. Some socio-anthropologic and organizational issues. Pediatrics 45: 868- 873 (1970).
- Weckworth, V.E. The conceptual model for reporting in Children and Youth projects. American Journal of Public Health 61: 78-89 (1971).
- _____. On evaluation: A tool or a tyranny - II. Systems Development Project Comment Series 9-11 (22), 1969.
- Weiss, R.S., & Rein, M. The evaluation of broad-aim programs: A cautionary case and a moral. Annals of the American Academy of Political and Social Science 385: 133-42 (1969). Reprinted in Caro, op. cit., 1971, pp. 287-296.
- Werner, E. Cumulative effect of perinatal complications and deprived environment on physical, intellectual and social development of pre-school children. Pediatrics 39: 490-505 (1967).
- Werner, E., Bierman, J.M., & French, F.E. The Children of Kauai: A longitudinal study from the perinatal period to age ten. Honolulu: University of Hawaii Press, 1971.
- White House Conference on Children, 1970. Profiles of Children. Washington, D.C.: Government Printing Office, 1970.
- Wholey, J.S. The absence of program evaluation as an obstacle to effective public expenditure policy: A case study of child health care programs. Unpublished paper prepared for the Joint Economic Committee (op. cit., 1969); circulated by The Urban Institute. Undated.
- Wholey, J.S., Scanlon, J.W., Duffy, H.G., et al. Federal Evaluation Policy: Analyzing the effects of public programs. Washington, D.C.: The Urban Institute, 1970.
- Wiener, G. Psychologic correlates of premature birth: A review. Journal of Nervous and Mental Disorders 134: 129-144 (1962).
- Wight, B.W. The control of child-environment interaction: A conceptual approach to accident occurrence. Pediatrics 44: Supplement (November, 1969), 799-805.
- Williams, W. Social Policy Research and Analysis: Experience in the Federal agencies. New York: American Elsevier, 1971.

- Williams, W., & Evans, J.W. The politics of evaluation: The case of Head Start. Annals of the American Journal of Political and Social Science 385: 118-32 (1969). Reprinted in Lyden & Miller, op. cit. (1972), pp. 374-392.
- Wolin, S.S. Politics and Vision: Continuity and innovation in Western political thought. Boston: Little, Brown and Company, 1960.
- Woolsey, S. Nutritional status of children. Unpublished paper prepared for Office of Management and Budget, October, 1971.
- Yerby, A.S. The disadvantaged and health care. American Journal of Public Health 56: 5-9 (1966).
- Yerushalmy, J. Neonatal mortality by order of birth and age of parents. American Journal of Hygiene 28: 244-270 (1938).
- Yerushalmy, J., Palmer, C.E., & Kramer, M. Studies in child-birth mortality, age and parity as factors in puerperal fatality. Public Health Reports 55: 1195-1220 (1940).
- Young, D.L. Analysis of the National School Lunch Programs and Child Nutrition Needs. Publication No. 705-51. Washington, D.C.: The Urban Institute, March, 1971.