United States General Accounting Office

GAO

Report to the Ranking Minority Member, Committee on Labor and Human Resources, U.S. Senate

January 1995

PROGRAM EVALUATION

Improving the Flow of Information to the Congress

GAO/PEMD-95-1

153296

GAO

United States
General Accounting Office
Washington, D.C. 20548

Program Evaluation and
Methodology Division

B-257897

January 30, 1995

The Honorable Edward M. Kennedy
Ranking Minority Member
Committee on Labor and Human Resources
United States Senate

Dear Senator Kennedy:

This report responds to your request that we assist the Committee in identifying and obtaining the information needed to oversee and evaluate the programs under its jurisdiction. We review the questions that the Committee is likely to find useful for evaluating programs of different types. We then explore the extent to which three programs collect information pertinent to these questions and report it to the Committee. Finally, we suggest a strategy that the Committee might use to identify evaluation questions of interest for a given program and to request information in advance from agency officials for later use in reauthorization or oversight hearings.

We will send copies of this report to the Secretaries of Education and Health and Human Services and to interested congressional committees. Copies will also be made available to others upon request.

If you have any questions or would like additional information, please call me at (202) 512-2900 or Robert L. York, Director of Program Evaluation in Human Services Areas, at (202) 512-5885. Major contributors are listed in appendix VI.

Sincerely yours,

Terry E. Hedrick
Assistant Comptroller General

Executive Summary

Purpose

Committees of the Congress need evaluative information to help them make decisions about the programs they oversee—information that tells them whether, and in what important respects, a program is working well or poorly. Concerned that the information it receives from administrative agencies is often insufficient, the Senate Committee on Labor and Human Resources asked GAO to (1) identify the kinds of information that would be useful for oversight and reauthorization review of the types of programs under its jurisdiction; (2) examine the extent to which agencies collect and report such information; and (3) propose a strategy the Committee could use to improve its access to agency information.

Background

The Senate Committee on Labor and Human Resources oversees most of the programs administered by the Departments of Education, Labor, and Health and Human Services (HHS). These departments fund a wide variety of programs, from basic medical research to education services for disadvantaged children. The Committee can request information in various ways, but it does not automatically receive evaluative information about each program as it comes up for legislative action.

GAO drew from the literature of program evaluation and consulted experts and Committee staff to identify categories of information and core questions within each category that would be useful for oversight and reauthorization review of various programs. Statistics, research, demonstration, service, and regulatory programs were included in the review. The information categories covered included descriptive information about the programs and evaluative information about program implementation, targeting, impact, side effects, and comparative advantage. GAO then examined documents and interviewed officials of three programs to ascertain which core questions were pertinent to each program, what information relevant to those questions was available, what information had been requested by and reported to the Committee, and how requests and responses had been communicated. The three programs, selected with the Committee's agreement, were

- the Comprehensive Child Development Program (CCDP), which aims to demonstrate that providing very young children and their families with educational, health, and social services contributes to child development and family self-sufficiency;
- the Community Health Centers (CHC) program, which provides clinical services to medically underserved areas or populations; and

GAO's Analysis

Committee staff accorded good, appropriately disaggregated descriptive information high priority and indicated that it would answer many of their questions about a program. Information about implementation is critical for oversight of new programs and provisions and for evaluation of programs whose activities must be implemented in accordance with professional standards if they are to succeed. The relevance of questions about program targeting, impact, side effects, and comparative advantage reflects features specific to each program and is best determined on a case-by-case basis.

The programs included in this study offer contrasting examples of information requested, collected, and received. The CCDP was designed with evaluation in mind and collected a wealth of descriptive information on clients and services, as well as data that could answer evaluative questions about program implementation and effects. Although the mandated evaluation report on the program was still in preparation, some emerging findings about CCDP had reached the Committee through various means in 1993 and early 1994. The CHC program collected information for program management purposes that could be organized to answer questions about service quality, program targeting, and other relevant issues. Agency staff used this information to answer specific questions from the Committee, but available information had not been synthesized and much of it remained unreported. The Chapter 1 ESEA program collected little management information from grantees but drew on data from a wide range of studies that were summarized in a mandated evaluation report in 1993.

In each of these cases, the Committee obtained the program information it requested. It asked for and received information on particular aspects of the CHC program. The Congress mandated studies on each of the other two programs, and the Committee received reports accordingly. However, the information the Committee had initially asked for—once received—did not always meet its current needs. For example, information on discrete aspects of the CHC program did not convey an overall picture of the program. Both the CCDP and the Chapter 1 ESEA evaluation mandates as expressed in the statute presented feasibility problems or design constraints that limited the usefulness of results. (For example, information on program impact was to be presented before new provisions could have been fully implemented.) In both cases, the policy issues had also changed somewhat by the time the report was due.

- the Chapter 1 Elementary and Secondary Education Act (ESEA) program of grants to states, which in turn fund school districts to provide compensatory education services in schools serving low-income students.

Drawing on the broad review of information needs and the case studies, GAO framed a strategy the Committee could use to request information from an agency in order to evaluate whether a program is adequately serving its intended purpose. Of course, the purpose itself may need to be reevaluated. However, such reevaluations draw on information from many sources, not primarily from the administering agency, and thus fall outside the scope of this report.

Results in Brief

GAO concluded that a brief list of descriptive and evaluative questions could capture the agency information most useful for program evaluation. Questions can be selected and adapted from this list to fit the program and the Committee's purpose in seeking information. Descriptive information is essential to the Committee and should be disaggregated to show how activities, settings, and clients vary within a program. For oversight purposes in the years between reauthorizations, it is important that the Committee receive information on progress in implementing new provisions and notification of new developments that may require future attention. Reauthorization decisions may call for systematic answers to evaluative questions about program implementation, targeting, impact, and side effects. When associated with policy review, such decisions may also call for comparison of the current program to various alternatives.

Each of the programs GAO studied collected a great deal of useful information. However, much of this information was not requested and did not reach the Committee. Information that was specifically requested did reach the Committee, but much of it was difficult to digest, too highly aggregated, or received too late to be useful. Communication between Committee and agency staff on information issues was limited and afforded little opportunity to build a shared understanding of the Committee's needs and how to meet them.

GAO concluded that obtaining useful information involves selecting pertinent descriptive and evaluative questions, explicitly requesting a response, and communicating with agency staff to ensure mutual understanding of what is needed.

GAO observed that contacts between Committee and agency staff about these programs consisted of a series of one-way communications (from the Committee to the agency or the reverse), with little opportunity for direct discussion. There appeared to be no current mechanism for adjusting requests in the interest of technical soundness or new priorities, or of adjusting reporting plans and formats to increase the timeliness and accessibility of the information reported. Communications between the executive and legislative branches on evaluative questions are policy-sensitive and follow formalized procedures. However, increased communication about information needs could be accommodated within these procedures.

Matters for Congressional Consideration

This report proposes a strategy for obtaining information to assist with program oversight and reauthorization review that the Committee may wish to adopt. The three components of this strategy are (1) selecting and adapting, from a core list, the descriptive and evaluative questions to be asked about a program in interim years and at reauthorization; (2) arranging explicitly to obtain timely oversight information in interim years as well as to receive results of evaluation studies at reauthorization; and (3) providing for increased communication with agency program and evaluation staff to help ensure that information needs are understood and that requests and reports are suitably framed and are adapted as needs evolve. This strategy can be adapted to take institutional realities into account. For example, in view of the many demands on its attention, the Committee might select future reauthorization questions for some programs and invite agencies to propose questions for others.

Agency Comments

Responsible officials from the Departments of Education and HHS provided written comments on a draft of this report. (See appendixes IV and V.) Officials from both agencies concurred with GAO's suggestion that the dialogue between the Committee and the agencies be strengthened at critical points in the evaluation process. While generally supporting GAO's approach, agency officials were concerned that it be consistent with the requirements of the Government Performance and Results Act (GPRA) and not constitute an added burden. GAO considers its approach to be compatible with GPRA and foresees little, if any, added burden on the annual reporting of program performance data. When the Committee requests information about other aspects of a program, such as side effects or comparative advantage, consultation should ensure that burden is taken into account as evaluations are planned.

Contents

Executive Summary		2
Chapter 1 Introduction		8
	Background	8
	Objectives, Scope, and Methodology	9
	Types of Programs the Committee Oversees	10
	Case Study Programs Reviewed	12
	Strengths and Limitations of Our Study	13
	Organization of This Report	14
Chapter 2 What Information Would Be Most Useful?		15
	Categories of Information and Related Questions	16
	Different Forms of Program Information and Why They Are Needed	23
	Summary Observations	24
Chapter 3 What Information Is Available? Three Case Studies		27
	Introduction	27
	Demonstration Program: CCDP	27
	Service Programs: CHC and Chapter 1 ESEA	31
	Communicating About Information Needs	37
	Concluding Observations	39
Chapter 4 A Strategy for Requesting Evaluation Information		41
	Identifying and Adapting Core Questions	41
	Arranging for Responses	45
	Communicating About Information Needs	47
	Feasibility of Our Approach	48
	Agency Comments and Our Response	49
Appendixes	Appendix I: Comprehensive Child Development Program	52
	Appendix II: Community Health Centers	58
	Appendix III: Chapter 1 Elementary and Secondary Education Act	64
	Appendix IV: Comments From the Department of Education	72
	Appendix V: Comments From the Department of Health and Human Services	74
	Appendix VI: Major Contributors to This Report	77

Bibliography 78

Related GAO Products 84

Tables

Table 2.1: Core Questions by Category of Information 26
Table 3.1: Core Questions Adapted to CHC and Chapter 1 ESEA Programs 32
Table 4.1: Guide to Identifying Needed Information 44
Table I.1: CCDP Evaluation Questions and Related Information 55
Table II.1: CHC Evaluation Questions and Related Information 61
Table III.1: Chapter 1 ESEA Evaluation Questions and Related Information 68

Abbreviations

ACYF	Administration on Children, Youth, and Families
AFDC	Aid to Families With Dependent Children
BCRR	Bureau's Common Reporting Requirements
BPHC	Bureau of Primary Health Care
CCDP	Comprehensive Child Development Program
CHC	Community Health Centers
ESEA	Elementary and Secondary Education Act
GPRA	Government Performance and Results Act
HHS	Department of Health and Human Services
HMO	Health maintenance organization
HRSA	Health Resources and Services Administration
NACHC	National Association of Community Health Centers
PHS	Public Health Service

Introduction

Congressional committees need information to help them make decisions about the programs they oversee—information that tells them whether, and in what important respects, a program is working well or poorly. Executive agencies provide evaluations, studies of the implementation of programs, and other reports that can contain information helpful to congressional committees. But authorizing committees may not automatically receive such information about an existing program as it comes up for reauthorization or oversight review and thus may receive insufficient program information.

The then-Chairman of the Senate Committee on Labor and Human Resources expressed concern that the Committee was not getting the evaluative information it needs to carry out its oversight responsibilities and asked us to ascertain what information would be most useful and how this information could be obtained from the responsible program agencies.

Background

This authorizing committee oversees programs administered by the Departments of Education, Labor, and Health and Human Services. These programs perform functions as diverse as funding education for disadvantaged children, providing medical care, conducting basic medical research, collecting data on American jobs, and enforcing occupational safety standards.

The Committee has used a variety of strategies to request or obtain information on these varied programs. For some programs, it has proposed (and the Congress has adopted) legislation that identifies specific questions to be answered in an evaluation study or annual report. For others, legislation has mandated an evaluation study or report of activities, but in very general terms. Another legislative strategy has been to direct an agency to set aside some portion of its appropriation for evaluation, yet permit the department to decide which programs to evaluate as well as which questions to ask. The Committee also obtains program information by asking questions as they arise or in connection with oversight and appropriations processes.

The Committee can also draw information from reports that all agencies must provide, such as appropriation request documents, the Chief Financial Officers Act report, and the Federal Managers' Financial Integrity Act report. In the future, the Congress will also receive reports on program performance and outcome measures under the Government Performance and Results Act. Yet despite these different ways of obtaining

information, the Committee finds that it does not necessarily get the information it needs, when it needs it, in a usable form.

Objectives, Scope, and Methodology

Objectives and Scope

The then-Chairman of the Senate Committee on Labor and Human Resources asked us to help the Committee learn more about how to evaluate the programs under its jurisdiction. First, the Committee requested that we examine the effectiveness of a specific evaluation provision, one permitting a set-aside of program funds for evaluation of public health programs. Our 1993 report found that these funds were not fully utilized and that evaluation results were not synthesized or regularly communicated to the Congress, suggesting that other approaches for getting information to the Congress were needed.[1] Second, the Committee asked us to identify information needed to evaluate the types of programs it oversees and to suggest how it might request such information from the agencies.

This report addresses the second task. Based on the letter of request and discussions with Committee staff, we identified three report objectives: (1) to identify the kinds of information that would be useful to the Committee for oversight and reauthorization review of the various types of programs under its jurisdiction, (2) to examine the extent to which the agencies currently collect and report such information, and (3) to propose a strategy the Committee could use to improve its access to useful information from the agencies.

Our objective was to help the Committee obtain agency information about the performance of existing programs, to assist them with decisions such as how to refocus or improve an existing program that meets an evident need. Of course, decisions about a program's future may involve fundamental policy questions—such as the question of whether program continuation is warranted in light of social and demographic changes, current budget conditions, and policy priorities. Information to resolve such questions comes to the Committee from a wide range of sources, not primarily from the administering agency. Since our concern was with how

[1]See Public Health Service: Evaluation Set-Aside Has Not Realized Its Potential to Inform the Congress (GAO/PEMD-93-13; Apr. 8, 1993).

to obtain information from administrative agencies, we touch on the fundamental policy questions only briefly. However, readers should bear in mind that broad-based policy evaluation often takes place prior to or in tandem with the evaluation of program performance.

Methodology

To identify the kinds of information that would be most useful, we (1) consulted the literature and experts familiar with the evaluation of public programs of various types, (2) asked Committee staff members what information they and Committee members would most like to have for oversight and reauthorization, (3) reviewed statutory and other requests for information about our case-study programs, and (4) compiled a list of core questions drawing from all of the foregoing. In this way, we identified categories of information that are useful for evaluation purposes as well as for ongoing oversight activities.

To learn how much of this information agencies collect and report, we narrowed our scope to conduct in-depth case studies of three programs selected after consultation with the Committee. (These programs and the selection criteria are described in more detail later in the report.) For each program, we reviewed documents and conducted interviews with agency staff to inventory what information was available to answer core questions pertinent to that program. We also ascertained from Committee and agency staff what information the Congress had requested about each program, what Committee and agency communication had occurred around these requests, and what the Committee had received. We did not assess the quality of the agencies' data.

To frame a strategy the Committee could use to request information, we drew upon our observations concerning major gaps in the information the Committee receives and shortcomings in communications between the Committee and program agencies. Responsible officials from the Departments of Education and HHS commented on our findings at exit conferences and provided written comments on a draft of this report. (See appendixes IV and V.) We conducted our work from July 1993 to October 1994 in accordance with generally accepted government auditing standards.

Types of Programs the Committee Oversees

Our first objective required that we classify the programs under the Committee's jurisdiction by program type, using a typology that had relevance for evaluation. Our review of public policy and program

evaluation literature found that a program's purpose is central to any consideration of how to evaluate its effectiveness or worth.[2] Five major program purposes—statistics, research, demonstration, service, and regulation—appeared relevant to the work of the Committee. We describe the five here as ideal types.

Statistics programs compile and analyze data and disseminate information, typically on a recurring basis. They develop indicators from data collected for administrative, enforcement, or statistical analysis purposes relevant to public policy issues such as health care, employment, or education. Programs of this type under the Committee's jurisdiction include the Bureau of Labor Statistics and the National Center for Education Statistics.

Research programs primarily aim to develop new knowledge or increase understanding of the subject studied. Research may be conducted in-house by agency personnel or by individuals, research institutions, or other organizations supported by grants or contracts. Examples of research programs are the National Institute on Drug Abuse and the Research and Demonstration Center program of the Office of Educational Research and Improvement.

Demonstration programs are defined here as those that aim to provide evidence of the feasibility or effectiveness of a new approach or practice. They may be undertaken to learn about the suitability of a particular approach under real-world conditions, to try out a variety of approaches and examine their effects, or to develop a new approach.[3] Examples are the Jobstart demonstration program and the Comprehensive Child Development Program.

Service programs provide services, directly or indirectly, to users or clients. Not only are there many different kinds of service (human, organizational, informational, financial), but they are delivered and funded through a variety of mechanisms, making service the most varied of the program types. Some programs deliver a specific service (such as student loans for postsecondary education) while others (such as the Community

[2]Some public policy literature classifies programs according to the "tool" through which they are carried out (for example, through a government agency, project grant, formula grant, loan, loan guarantee, or regulation) or by their maturity (whether the program is new or established). Program purpose is the basis for our typology, but the other dimensions are discussed where relevant.

[3]The term "demonstration" may also refer to activities whose purpose is simply to provide an example of an approach. In our view, such purely illustrative demonstrations are most appropriately classified (for these purposes) as service programs—since they mainly do not result in evaluative conclusions about an approach.

Services Block Grant program) support a broad range of services that may vary across providers or locations. Government agencies may make grants or loans directly to recipient individuals or organizations, to service providers, or to states (which in turn make subgrants to providers).

Regulatory programs are intended to protect people from harm or from a violation of their rights by influencing the behavior of individuals or organizations. Some regulations prescribe actions (such as requiring employers to follow certain workplace safety practices) and others prohibit actions (such as employer discrimination against employees on the basis of age). Federal agencies (directly or through the states) adopt rules or standards of behavior and compel or coax public compliance with the regulation. Examples are the Occupational Safety and Health Administration and the Department of Education's Office of Civil Rights programs.

We reviewed programs overseen by the Committee and found that we could classify them according to these types.[4] To do so, we examined the programs' actual functions, not simply the title or label given to a program by enabling legislation. (For example, if a "demonstration" program focused on providing services rather than on gathering evidence about what was demonstrated, we treated it as a service program.) In addition, we observed that some programs have a secondary purpose that complements their primary purpose, such as a statistics program that is intended not only to produce data, but also to perform the service of disseminating data to clients. Such programs would need to be evaluated in terms of each relevant purpose.

Case Study Programs Reviewed

To address our second objective and assess the extent to which potentially useful descriptive and evaluative information was collected and reported by federal agencies, we conducted case studies of three programs the Committee oversees. The following three programs were selected, in consultation with the Committee staff, from a list of programs due to be reauthorized:

- the Comprehensive Child Development Program, which aims to demonstrate the effectiveness of various approaches to providing very young children and their families with educational, health, and social services;

[4]Our typology is comprehensive for all programs under this Committee's jurisdiction. It does not cover other possible program types, such as cash assistance programs, entitlement programs (such as Social Security), or tax-expenditure programs.

- the Community Health Centers program, which funds centers that provide clinical services to medically underserved areas or populations; and
- the Chapter 1 Elementary and Secondary Education Act program of grants to states, which in turn fund school districts to provide compensatory education services in schools serving low-income students.

The Chapter 1 ESEA program is administered through the Department of Education, and the other two programs are administered through HHS. (For program details, see appendixes I-III.)

This portfolio of one demonstration and two service programs covers only two of our five program types. However, our three cases offer the following advantages:

- All three areas—health, education, and early childhood development—are high on the current policy agenda.
- Information about a demonstration program and its evaluation is particularly important, since a demonstration (such as CCDP) provides the Congress with an opportunity to learn the effects of a proposed new program on a small scale before committing to major implementation.
- Service programs constitute much of the Committee's workload and funding authority. Our cases represent two major variants—one provides federal grants to the states, which in turn fund local providers of services (Chapter 1 ESEA), and the other provides grants from the federal government directly to service providers (CHC).

Strengths and Limitations of Our Study

The major strength of our study is that, in identifying this Committee's evaluation information needs, we considered the structure of the legislative process, the nature of the programs, and the agencies' program administration practices. Both the Congress and federal agencies could find this perspective useful in developing both better information and a better match between evaluative needs and the information provided.

Another strength of the study is our examination of the communication process between this Committee and the three agencies it oversees, including how that process affected what information was reported and when it was received. Our report also provides detailed information on the three programs, information for which there is both an immediate and a longer term need.

A limitation of this report is that the three programs, while illustrative, are not representative of all programs. The case studies exemplify the importance of, and barriers to, gathering (or not getting) program information but do not, of course, reflect all possible types of problems. In addition, observations based on the study of a particular Senate authorizing committee at a particular point in time may not reflect the information needs and procedures of other Senate or House authorizing committees or of appropriations committees. However, we are confident that the general strategy we propose could be readily adapted to cover other policy mechanisms (such as tax incentives) and evaluation questions (such as whether there is still a demonstrable need for the program) that we did not specifically address.

Organization of This Report

Chapter 2 describes the kinds of information we found to be of most potential use to the Committee, as well as how the core questions to be asked vary by program type. Chapter 3 summarizes the information available for each of our three case programs and examines congressional requests for information and how agencies responded to them. Chapter 4 proposes a strategy the Committee staff could use to obtain useful evaluative information in the future.

Appendixes I-III contain detailed information on the three programs reviewed in chapter 3. Each appendix includes a brief description of the program and a table listing evaluation questions, specific information needed to answer those questions, and an assessment of whether the needed information was available. Appendixes IV and V reproduce the comments from the Departments of Education and HHS on a draft of this report.

What Information Would Be Most Useful?

Our first task was to determine what kinds of information would be most useful to the Committee as it evaluates various types of programs for possible reauthorization and oversees their progress during interim years. Our approach to this task included reviewing the literature on public program evaluation and congressional oversight and interviewing experts familiar with the evaluation questions associated with different types of programs as well as congressional requests for information about such programs.[1] We discussed the results of this inquiry and their own views with senior members of the Committee staff with responsibilities in each of several program areas and obtained their assessments of the kinds of information they and Committee members would most like to receive.

From this approach, we identified priority categories of evaluative information, compiled a list of core questions connected with each category, and noted the kinds of data or other material that would be needed to respond to these questions. We examined the relevance of the various core questions to programs of different types. Finally, we identified different forms of information that are useful for different purposes, such as oversight, reauthorization, and policy change.

Our inquiry focused primarily on agency-provided information that could help the Committee evaluate how well an existing program is working. Such information is helpful in deciding whether a program is adequately serving its intended purpose and if so, what level of funding is appropriate. Of course, the purpose itself may need to be reevaluated: changing conditions may have rendered it irrelevant or decreased the need for the program. Because such reevaluations draw on information from outside the program agency, they fall outside the scope of this report and we do not discuss them specifically. However, guidelines to identifying the information needed for such studies can be found in the general literature of evaluation and in other GAO reports.[2] The general strategy we outline in chapter 4 should assist the Committee in obtaining information concerning the need for a program, as well as the other categories of information outlined in the remainder of this chapter.

[1]We interviewed program specialists from the four congressional support agencies: Congressional Research Service, Office of Technology Assessment, Congressional Budget Office, and GAO.

[2]For a discussion of criteria for evaluating the need for a program, see Children's Programs: A Comparative Evaluation Framework and Five Illustrations (GAO/PEMD-88-28BR; Aug. 31, 1988).

Categories of Information and Related Questions

The main congressional evaluation questions we derived can be broadly categorized as seeking **descriptive** information about the program, information that can be used to evaluate its **implementation**, and information about various **effects**. These three broad categories of information are generally applicable to all five of the program types we identified. However, the relative priority of different categories and questions varied somewhat from type to type as did the particular manner in which our core questions were phrased. We describe our information categories and core questions next.

Program Description

Descriptive information tells what the program is and what it does. Descriptive questions include the following: What activities does the program support? Toward what end or purpose? Who performs these activities? How extensive and costly are the activities, and whom do they reach? Are conditions, activities, purposes, and clients fairly similar throughout the program, or is there substantial variation across program components, providers, or subgroups of clients?

The descriptive information that such questions elicit is not itself evaluative and receives relatively little emphasis in the program evaluation literature. It is important for evaluation, nonetheless, and was accorded high priority by the Committee staff we interviewed. Descriptive narration and statistics provide the foundation for identifying key evaluation questions and interpreting the evaluative information concerning the program. Better descriptive information would answer many of the Committee's questions about the program and thus would enable the Committee to focus its oversight and evaluation efforts more effectively.

Committee staff indicated that they typically do not get sufficient descriptive information to understand the variety of conditions under which a program actually operates and how federal funds are actually being used. (General summaries of activities and programwide totals of actions performed or clients served, such as accompany appropriation requests, do not meet this need.) Nor do they typically receive site-specific observational data to convey the "flavor" of a program as participants see it (for example, to help them understand how a student's experiences in federally supported bilingual education classes relate to the rest of his or her school day). Such information, they told us, is essential for understanding how well a program is working and thus is relevant to the Committee even when site-level conditions are a state or local responsibility.

Our review and interviews suggested that descriptive information is useful for every type of program, be it concerned with statistics, research, demonstration, service, or regulation. The nature and amount of information needed vary with such program features as internal diversity and contact with the public. For example, the typical summary description may provide a sufficient basis for understanding a statistics program or a demonstration that performs a uniform set of activities in a uniform way in one or a few sites. Much more information, appropriately disaggregated, is needed to understand a demonstration, service, or regulatory program that operates through diverse providers who perform different mixes of activities under diverse conditions (for example, in schools, clinics, or workplaces ranging from the very large and urban to the very small and rural).

Implementation

Our second category concerns information about program implementation—specifically, about how and to what extent activities have been implemented as intended and whether they are targeted to appropriate populations or problems.

Extent and Nature of Implementation

Questions about the extent and nature of program implementation are concerned with procedural issues—that is, with how program activities are carried out. Relevant questions include the following: (1) Are mandated or authorized activities actually being carried out? (2) Are the activities in accordance with the purpose of the law and implementing regulations? (3) Do they conform to the intended program model or to professional standards of practice, if applicable?[3] (4) Are program resources efficiently managed and expended?[4]

These implementation questions typically arise in the years immediately after a new program or provision is authorized, as well as at reauthorization. With respect to new provisions, our respondents were interested in learning whether acceptable progress had been made in putting new activities into practice, whether significant feasibility or management problems had arisen, and whether program modifications were needed as a result. Information obtained through the administering

[3]This and the preceding question refer to what is sometimes called "program fidelity" in the evaluation literature.

[4]Program evaluations and reauthorization discussions typically do not focus on management issues unless there is evidence of problems with the program in question. Our review accordingly did not focus on such issues or the information needed to address them. However, we consider it useful for the Committee to ask whether there is evidence of such problems and, if so, to request follow-up information.

agency's standard reporting and monitoring procedures may be sufficient to answer these interim-year questions and to identify feasibility or management problems that arise in connection with ongoing activities. For reauthorization, interest focused on questions concerning conformance to program purposes, model specifications, or professional standards.

Descriptive data obtained through routine reporting, if designed with this purpose in mind and known to be sufficiently reliable, may provide a sufficient basis for comparing program activities against program requirements.[5] However, special data collection—obtained through program monitoring or through external studies—may be necessary to ensure the objectivity and detail required to explore issues of program quality.

Our review indicated that most of the implementation questions we have discussed are pertinent to all of the types of programs we considered. However, there are some differences with respect to the question of conformance to professional standards. This question not only applies but also is critical to the evaluation of two of the types of programs we reviewed: statistics and research. Well-established professional standards apply to these types of programs, and procedures and products that do not meet such standards will not be considered credible.[6] Applying these standards, an evaluation of the National Center for Education Statistics should ask whether the education indicators generated were free from bias and technically sound and whether data collection methods, analysis methods, and limitations of the data were fully disclosed. Expert judgments commonly are used to provide the needed evaluative information.

Obtaining information on the extent to which prescribed procedures were followed is also especially important for demonstration projects. If a prescribed approach was to be demonstrated, the Committee will want to know whether the approach was implemented as planned and, if not, how

[5]Client counts and other monitoring data, dependent as they are on the limited resources that program offices or grantees typically can devote to data collection, may well be imprecise. The cost of achieving precision is considerable, and often outweighs the benefits. In such instances, the agency should make a systematic estimate of the extent of uncertainty in their numbers and then present this estimate along with the numbers themselves.

[6]Criteria to be met by federal statistical agencies have been set forth by the National Research Council. See Margaret E. Martin and Miron L. Straf (eds.), Principles and Practices for a Federal Statistical Agency (Washington, D.C: National Academy Press, 1992). The Council's report emphasizes that statistical agencies' work will have the necessary credibility only if it is impartial, technically sound, and fully described. Similar professional criteria apply to research.

it was modified and why modifications were deemed necessary. Similarly, determining how activities and arrangements differ across projects is necessary in order to assess those demonstration programs designed to encourage and compare alternate approaches.

For regulatory and service programs, the applicability of models or standards and the importance of evaluating conformance to them must be examined on a case-by-case basis. Some regulatory and service programs emphasize the importance of following specific procedures (such as medical procedures or procedures for gathering evidence of compliance with regulatory requirements), while other programs involve activities for which standards are not available or have not been specified.[7] (See our case studies of service programs in chapter 3.)

Targeting

Targeting has to do with the objects (problems or populations) at which a program's activities are directed and the coverage that these activities achieve. Many programs under the Committee's jurisdiction are targeted to particular populations or problems (such as migrant students or unsound management of pension plans). Others are directed to a broad area (such as research on infectious diseases), from which the agency is expected to select high-priority problems or clients (such as a particular disease that endangers children or the elderly). In either case, it is useful for the Committee to know what kind of coverage a targeted program has achieved. Relevant questions include the following: (1) Have program efforts focused on appropriate problems? (2) Has the program reached the appropriate people or organizations? (3) Do current resources and targeting practices leave significant needs unmet (that is, priority problems unaddressed or priority clients unserved)?

Assessing a program's targeting success typically requires not only good measures of program activities and data on the clients served, but also information about the size and distribution of the eligible population (for target groups) or problem area as a whole. Thus, this assessment may require a special data collection activity or access to sources of data outside the program (such as the decennial census).

The question of whether appropriate problems are targeted is critical to statistics programs (which should address policy-relevant issues) and research programs (which should address important gaps in knowledge). For other types of programs, the importance of targeting and the particular

[7]Our focus here is on standards for conducting the service or regulatory activities central to the purpose of the program. Federally funded programs of every type are expected to meet financial and certain other management standards.

questions to be asked must be determined on a case-by-case basis. Questions for programs of different types are illustrated below.

- **Statistics**: Are Bureau of Labor Statistics indicators relevant to current issues of public policy? To what extent have they reached key federal, state, and private sector policymakers? Are there important issues or populations on which labor statistics are not being collected?
- **Service**: What proportion of the students served by the migrant education program are currently migrant (that is, have moved within the past 5 years)? What proportion of recent migrants are not served?
- **Regulatory**: Are Occupational Safety and Health Administration enforcement efforts targeted to the safety problems that pose the greatest dangers or that endanger the greatest number of people?

Program Effects

We found it useful to distinguish three different aspects of program effects: (1) whether the program is achieving its intended purposes or outcomes (impact), (2) whether it has other important effects that relate to congressional concerns (side effects), and (3) how it compares with alternative strategies for achieving the same ends (comparative advantage). We address each of these aspects below.

Impact

Impact questions center on whether program activities actually resulted in the improvements the program was designed to produce, as well as on what progress was made toward achieving the program's goals. Where programs or demonstrations are designed to produce changes **as a result of** program activities, assessing a program's impact is often central to reauthorization deliberations. For example, a job training program is expected to show that participation in the program led to significantly higher income or more stable employment. Where impact was expected, our respondents were interested in learning, for reauthorization purposes, about (1) the aggregate impact of the program; (2) how impact or outcomes varied across participants and approaches; and (3) how impact varied across providers—specifically, whether the program was supporting providers whose performance was consistently weak.

Good evaluative information about program effects is difficult to obtain. Each of the tasks involved—measuring outcomes, ensuring the consistency and quality of data collected at various sites, establishing the causal connection between outcomes and program activities, and separating out the influence of extraneous factors—raises formidable technical or logistical problems. Thus, evaluating program impact

generally requires a planned study and, frequently, considerable time and expense. Program features affect the relative difficulty of getting good impact information. The more varied the program activities, and the less direct the connection between the provider and the federal agency, the greater the likely difficulty of getting comparable, reliable data concerning clients and services. For example, a federal agency whose own employees deliver a specified service is likely to be able to obtain impact data more easily than one that administers grants that states then pass on to a variety of local entities to be used in any of several ways. Also, due to the absence of a contrasting comparison group, it is practically impossible to estimate the impact of a long-standing program that covers all who are eligible.

It is critical to obtain impact information from demonstration programs intended to test whether an approach (or any of several approaches) can obtain results, as well as from regulatory or service programs that are intended to produce specified kinds of outcomes or changes. However, impact questions are not commonly directed at statistics programs (which are designed to produce a product rather than a change). Nor are they directed at programs that regulate or offer a service when it is considered certain that the desired outcomes will occur as a result of program activities (for example, that a vaccine, properly administered, will produce immunization or that safety regulations, properly enforced, will reduce workplace injuries). Information showing that activities have been soundly implemented and have reached a high proportion of the relevant client population may be sufficient to justify a decision to continue such programs.

The purpose of research programs is to produce new knowledge, either basic or applied. The impact of basic research programs depends on whether the projects, in the aggregate, have added to the knowledge in their fields by eliminating or confirming hypotheses (for example, hypotheses concerning brain chemistry or human perception and cognition). Applied research, on the other hand, is expected to result in the development of new or improved products (such as better medications or new approaches to the teaching of reading). If it achieves its intended immediate impact, either type of research program may ultimately have a much broader social impact (such as a reduction in the incidence of alcoholism or a rise in educational achievement levels). However, it would be inappropriate to evaluate either type of program primarily in terms of such ultimate impacts. The influence of basic research on either product development or practice is hard to trace and may not be evident for many

years, and products or practices may fail to gain widespread use for reasons that are beyond their developers' control.[8]

Side Effects

Public programs of any type may have important (and sometimes unforeseen) effects beyond those they were intended to produce. These side effects, which were of interest to our Committee respondents, should be noted in any comprehensive evaluation. For example, a program might have unforeseen effects—either positive or negative—on the problem or clients it was designed to address. Or, a new program might have an effect on other programs aimed at a similar problem or population.

Information to identify unforeseeable side effects is hard to plan for except by encouraging reflective observation and maintaining open channels of communication between program administrators and clients. However, some kinds of side effects can be foreseen—indeed, possible side effects may well come up in debate when a program is authorized—and attention to them built into the program's data collection plans. For example, a program might arrange from the outset to collect data that would enable it to determine whether its procedures were unintentionally resulting in unequal treatment of similarly situated individuals. Or, a service program that refers clients to another program might arrange to monitor the level of demand placed on the other program in order to ascertain whether meeting this demand was impeding that program's ability to meet its own goals.

Comparative Advantage

Finally, evaluation may focus on the comparative advantage of continuing the current program. Typical comparative advantage questions include the following: (1) are the effects gained through the program worth its financial and other costs? and (2) taking both costs and effects into account, is the current program superior to alternative strategies for achieving the same goals? Options against which the current program might be compared could include:

- discontinuing the program;
- consolidating it with other federal programs that serve a similar purpose;
- utilizing a different type of program (such as tax incentives rather than regulation) to address that purpose;
- transferring responsibility for program decisions to the state or local level of government (for example, by replacing a federal categorical service program with block grants to support locally determined services); or

[8]For a comprehensive recent treatment of the subject, see Barry Bozeman and Julia Melkers (eds.), Evaluating R&D Impacts: Methods and Practice (Boston: Kluwer Academic, 1993).

- transferring the function to the private sector.

Comparative advantage questions typically arise when programs are up for reauthorization. Our Committee respondents indicated that such questions were of special interest in the current climate of budget pressure (which may force choices) and of major policy change (which may affect all programs in a particular policy area).

Comparison of the current program to alternatives raises special challenges for information collection. It requires not only good, comprehensive data on the current program, but also truly comparable data concerning alternatives or policy options. Simulation data or other forms of data based on projection may be needed when considering options for which there is no directly relevant experience. Because of their complexity, comparative studies are likely to require considerable lead time, careful planning, and special arrangements for implementation.

Different Forms of Program Information and Why They Are Needed

We observed from our interviews that program information serves several different functions for the Committee—functions that call for the receipt of this information at different times and in different forms. The primary functions of program information are to (1) signal the need for attention to a program in the interim years between reauthorizations, (2) guide reauthorization decisions, and (3) assist the Committee in deciding whether a major policy change should be recommended. We refer to the corresponding forms of information as (1) notification, (2) answers, and (3) evidence of comparative advantage.

Notification That Attention Is Needed

Our Committee respondents emphasized that it was very important that they be kept informed of significant developments on an ongoing basis and particularly that they receive early notification when a problem or the need for change in a program becomes evident (for example, notification that a new provision is having unwelcome side effects or that shifting needs have rendered program targeting obsolete). Such notification helps the Committee focus its efforts on the programs that need attention, avoid unpleasant surprises, and recognize successful program sites or practices.

Our interviews with Committee and agency staff revealed that developments that may merit notification come to their attention in a number of ways, often as a by-product of program management. Initial clues to such developments can emerge from observations made during

site visits, from the review of applications or of program performance measures, from concerns that site officials raise in meetings, from conversations with clients, or from media coverage of a local program. Since a single clue may reflect erroneous information or represent an isolated instance, other evidence is needed to verify the existence and importance of the development suggested by the clue. The program managers we interviewed could draw on a variety of resources—networks and data sources, as well as their own expert knowledge of the program—to make such a verification and to judge whether there was sufficient evidence to justify notice to the Committee.

The varied evidence that supports managers' judgment that something worthy of notice has occurred is not generally thought of as a form of evaluation information. This may be because the evidence on which notice is based is rarely sufficient to support evaluative conclusions about a program. However, this form of evidence is sufficient to identify evaluation questions to which the Committee will want answers in the future. We therefore inquired about information and notification concerning current developments in our case studies, and we report our findings in chapter 3.

Answers to Guide Reauthorization Decisions

At times when major reauthorization decisions are being made, the Committee needs answers concerning evaluation questions such as those just discussed. Answers arise out of systematic evaluation studies. These studies must be planned and budgeted for in advance, information relevant to the questions the Committee will face must be gathered, and findings must be reported in time to inform deliberation.

Comparative Advantage Information for Policy Decisions

At times, the program under review concerns a policy area that is being reexamined. In this event, the Committee may need information that evaluates the current program against alternative strategies (including dropping the program if it cannot be justified in terms of cost-effectiveness or need). As previously noted, such comparative studies require data beyond that which the administering agency can provide and raise complex analytic (as well as policy) issues. Thus, these studies should be planned and arranged for well in advance.

Summary Observations

Our interviews and reading of the literature led us to three conclusions concerning what information would be useful.

1. Descriptive information is more useful to the Committee than has generally been realized. Such information should be sufficiently specific, detailed, and disaggregated to convey an in-depth understanding of the program.

2. In the interim years between reauthorizations, it is useful to the Committee to obtain information on progress and problems in implementing new provisions, as well as to be notified of developments that raise new evaluation questions.

3. Reauthorization decisions (including those associated with major changes of policy) call for information that provides answers to evaluation questions concerning program implementation and effects and the comparative advantage of continuing the program or pursuing other options. Advance planning is generally needed to identify questions relevant to each program and collect the data needed to answer them.

We further observed that the questions that it might be useful to ask, expressed in general form, comprise a relatively short list. We have listed these questions—grouped according to whether they concern program description, implementation, targeting, impact, side effects, or comparative advantage—in table 2.1. For any given program, some of these questions (but probably not all) will be relevant. Knowledge of program type is of some help in narrowing the list, but features specific to the program (such as expectations concerning targeting) must also be taken into account.

Table 2.1: Core Questions by Category of Information

Category of information	Core question
Description	Overall, what activities are conducted? By whom? How extensive and costly are the activities, and whom do they reach?
	If conditions, activities, and purposes are not uniform throughout the program, in what significant respects do they vary across program components, providers, or subgroups of clients?
Implementation	What progress has been made in implementing new provisions?[a]
	Have feasibility or management problems become evident?[a]
	If activities and products are expected to conform to professional standards or to program specifications, have they done so?
Targeting	Have program activities or products focused on appropriate issues or problems?
	To what extent have they reached the appropriate people or organizations?
	Do current targeting practices leave significant needs unmet (problems not addressed, clients not reached)?
Impact	Overall, has the program led to improvements consistent with its purpose?
	If impact has not been uniform, how has it varied across program components, approaches, providers, or client subgroups?
	Are there components or providers that consistently have failed to show an impact?
Side effects	Have program activities had important positive or negative side effects, either for program participants or outside the program?[a]
Comparative advantage	Is this program's strategy more effective in relation to its costs than others that serve the same purpose?

[a]Information important for interim years.

We used these core questions to categorize and summarize the information that was available from the three programs that we used as case studies, as reported in chapter 3 and in the case study appendixes. In chapter 4, we suggest how the Committee might draw from this list to request interim-year information and to frame evaluation questions for reauthorization review of programs of different types and features.

What Information Is Available? Three Case Studies

Introduction

The second objective of this study was to determine to what extent information is actually available to answer oversight and reauthorization questions pertinent to various programs. To address this objective, we used one demonstration program and two contrasting service programs as case studies.

We identified core questions (described in chapter 2) pertinent to each program based on our interviews and our review of enabling legislation and congressional requests. Next, we reviewed all the information available on the programs (from site reports and management information systems to formal studies) to determine what information was available on core evaluation questions (as adapted to fit each program), how much of this information made its way to the Committee staff, and when and in what form it had reached the staff. We also interviewed Committee and agency staff to learn how adequately the Committee members and staff felt that they had been informed, and what additional communication occurred as the agency sought to interpret and respond to the Committee's requests.

Our three cases offer richly contrasting examples of information requested, available and reported. The Comprehensive Child Development Program is a new demonstration program whose design included explicit provision for the collection of descriptive and evaluative information. Chapter 1 of the Elementary and Secondary Education Act is a long-standing education service program that collects minimal data from grantees but has been studied extensively through several mandated evaluations, the most recent of them reported to the Congress in 1993. The Community Health Centers program maintains extensive data on program operations but has not been asked for, nor has it provided the Committee with, comprehensive evaluation or program reports for years. Taken together, the three cases offer considerable insight into what information is collected and how it reaches (or fails to reach) the Committee.

Demonstration Program: CCDP

Core Questions About CCDP

The Comprehensive Child Development Program, which originated in 1988, was designed to demonstrate whether providing comprehensive and continuous services to very young low-income children, their parents, and

family members enhances the child's well-being as well as his or her physical, emotional, and intellectual development.[1] The various sites funded through the program offer certain services in common—parent education and training, case management, early childhood development intervention, health care, child care, and family and child needs assessment. However, sites vary considerably in the way they deliver these services and in the specific objectives emphasized. (For example, some sites focus on helping clients with day-to-day basic survival, while others also provide preventive counseling.)

The descriptive, implementation, and effects questions for oversight and reauthorization of the CCDP program are as follows: What is the basic model or approach tested here? How are the centers different or similar in terms of the local population, geographic location, and service provision? What are some of the start-up and implementation difficulties the centers faced? What are the principal program impacts? Have there been beneficial or detrimental side effects to this program or any aspect of it?[2]

Available Information Relevant to These Questions

On the whole, we found that CCDP was designed (through legislation and administration) to collect the categories of information critical to the evaluation of this demonstration program, and it has done so. The Administration for Children, Youth, and Families (ACYF), which administers the program, has descriptive and evaluative information to answer the core questions we identified.

Many sources provide the agency with useful information about the CCDP. Between the program's management information system and observational reports, the agency has the raw material from which to construct both quantitative and qualitative descriptions of the CCDP approach as it is demonstrated across the centers, as well as to measure CCDP's impact and document other effects. (See table I.1 for more details on the data available.)

Description

ACYF has rich descriptive information on program operations, clients, and settings for the approximately 34 centers. The agency's information system includes standardized information from each center on client needs as well as on what services are offered and how often they are provided. Information is also collected on client and community characteristics—

[1]Although the enabling legislation did not refer to CCDP as a demonstration program, the agency has collected implementation and impact data to meet the information needs of a demonstration program.

[2]See appendix I for more information on the program and what specific information is needed.

whether the center is located in a rural or urban area and what the center's philosophy and staffing patterns are—through center-written progress reports, program-staff site visits, and ethnographers' reports. ACYF thus has information that can communicate a "feel" for the challenges providers face in delivering a comprehensive set of services and how clients see the program as affecting their lives.

Implementation

We found much useful information on CCDP implementation processes and feasibility issues, as well as evaluative information on whether the program is being implemented as intended. ACYF has qualitative and quantitative information that is useful for both compliance assessment and service improvement. The ethnographers' reports from the first year of the program included specific information on start-up problems. They also described how site officials completed such necessary steps as setting up interagency agreements for service delivery and how grantees' approaches to service delivery were adapted as programs took shape. The grant application, as well as the ethnographer, site visit, and progress reports, provided additional insights into planned implementation.

Impact

The agency has data on participating children's developmental progress based on the results of standardized tests, as well as on whether parents gained parenting skills. Measures of short-term changes in parental income and employment status, as well as of general parental well-being and level of education, are also available. The agency has gathered similar information on a nonparticipating control group of families to compare with clients to determine whether any observed changes in the program families were a consequence of the demonstration program. Anecdotal information from participating family members (in progress or ethnographer reports) about how the program has improved their lives has also been collected. These data should be sufficient to support conclusions concerning the program's impact on low-income families and children.

Side Effects and Comparative Advantage

The agency has diverse sources of information that would capture CCDP's side effects (intended or not) on the local community and its social services agencies. CCDP evaluation studies have focused on this program alone and were not designed to compare CCDP with similar or different programs. A study comparing programs or models would need CCDP program data as well as information on other programs, and it would need to be arranged for separately. In addition, cross-site comparisons of effectiveness may be difficult to make since CCDP centers differ in many ways. If the Committee wants a comparison of the CCDP approach with

another, or an assessment of which CCDP features are most needed, it should request such a study.

What the Congress Requested and What It Received

Enabling legislation for the CCDP mandated a study that was to report on results of the program by October 1, 1993. The legislation set out evaluation objectives for a study (describe service delivery mechanisms and assess project impacts), specified that project impacts be measured with control group comparisons, and provided for an evaluation of program successes, but allowed the agency to select its evaluation design.

In response to this mandate, the agency provided the Committee with a report on program impacts on families and communities, an analysis of program implementation, and a description of families and projects.[3] It reported short-term effects on children, families, and mothers in terms of their education, income, health, and parenting skills. Because it was due before the demonstration was complete, the report provided only interim findings.

This mandated report had both strengths and limitations. It provided needed outcome conclusions, useful feasibility information, and a separate, readable executive summary. Programwide impact results (and answers to feasibility and process evaluation questions) were reported succinctly and clearly, but disaggregated statistics would have shown the Committee the range of performance and reflected possibly important different features of the CCDP model. The feasibility information would have been useful during the deliberations concerning reauthorization of the Head Start Program and its inclusion of CCDP-like services, but the report came out in May 1994 just as final action on the legislation was occurring.

In the interim, the Committee obtained program information through three other means (apart from normal ad hoc information requests). In December 1991, the program agency sent the Committee copies of a report it had prepared to inform the public and other agencies about CCDP. Although intended for another audience, the report contained information on program operations and differences among sites and also evaluative information on program feasibility of use to the Committee in its CCDP oversight and Head Start reauthorization activities.

[3]ACYF, "Comprehensive Child Development Program—A National Family Support Demonstration," interim report to the Congress, May 6, 1994.

In addition, the agency provided information on CCDP to a special commission (consisting of congressional committee staff—including representatives from this Committee—as well as agency officials and outside experts) formed to make recommendations concerning the Head Start Program. The CCDP program officer provided information to assist the commission in understanding the possible effects of extending the Head Start Program to include CCDP-like services to very young children. Finally, a member of the Committee staff requested and obtained current data on program impact—data more recent than the figures to be summarized in the CCDP interim report—from child development center directors.

ACYF plans to provide a final report to the Congress in 1996. It will be based on the entire 5 years of experience with the program and is expected to provide an assessment of the overall impact of CCDP.

Service Programs: CHC and Chapter 1 ESEA

Core Questions About These Programs

The first of our service programs, the CHC program, funds medical centers to provide outpatient care (and supporting services such as patient transportation) to populations that otherwise would lack access to such health care. Some community health centers serve remote rural areas that lack medical facilities, and others serve urban populations that do not have access to care through other providers. (For a description of the CHC program, see appendix II.) The Chapter 1 ESEA program provides financial assistance (through grants to states) to local education agencies to meet the special needs of educationally deprived children who live in areas with high concentrations of low-income families. Student demographics, educational achievement levels, and the services delivered vary considerably from school to school. For example, some schools assist low-performing students within the regular classroom, while others pull these students out to a special setting for assistance. (For a description of the Chapter 1 ESEA program, see appendix III.)

Table 3.1 summarizes the core questions adapted to fit the type and nature of each program. The two programs have certain questions in common. Both operate in diverse settings, so it is important to ask how services and conditions vary as well as to secure a good summary description. Both are

targeted to needy populations, so it is relevant to ask what proportion of eligible clients were served and whether significant numbers of eligible clients remained unserved. (For Chapter 1 ESEA, attention has focused particularly on eligible students in the neediest schools—that is, schools in which a high percentage of students come from low-income families.)

Table 3.1: Core Questions Adapted to CHC and Chapter 1 ESEA Programs

Category of information	Questions for CHC	Questions for Chapter 1 ESEA
Description	Overall, what services are provided and to whom? How do services, center locations, and mix of clients vary from site to site?	Overall, what services are provided and to whom? How do services and school and student characteristics vary from site to site?
Implementation	Do clinical practices meet professional quality standards?	
Targeting	What fraction of those who would otherwise lack access to care does the CHC program serve? How does this fraction vary across grantees? What significant areas, health problems, or populations are not covered?	What fraction of the neediest schools and students participate? How does this fraction vary across sites? How many sites lack services?
Impact	Has the long-term health status of CHC patients improved as a result of program services?	In the aggregate, has disadvantaged students' achievement improved? Has individual participants' achievement? How does impact differ across approaches and across schools? Do some schools or approaches consistently fail to produce gains?
Side effects	What are the effects on other health care providers (if any) in the CHC communities?	Does the program fragment responsibility for ensuring that students' progress improves?
Comparative advantage		What are the comparative advantages of shifting to schoolwide, rather than individually targeted, approaches?

Beyond the common ground delineated in table 3.1, the two programs are dissimilar. For the CHC program, the question of whether the clinical services provided meet current standards of good medical practice is fundamental. In terms of program effect, the simple provision of quality clinical services is an important program outcome. The broader medical

research question of whether these services influence a patient's health status in the long run is pertinent to any health care service system (not just this program) and thus needs to be answered in a broader context. The Chapter 1 ESEA program, by contrast, focuses on questions of impact, such as whether the provided services produced increases in student achievement. Defining and evaluating the quality of implementation of Chapter 1 ESEA services is a state and local responsibility. However, the Congress has an interest in learning whether some of the general approaches used are more advantageous than others.

Available Information Relevant to These Questions

The agencies that administer the service programs we reviewed collect a considerable amount of information potentially relevant to the foregoing questions. HHS' Health Resources and Services Administration (HRSA) receives narrative information in center applications; collects data on CHC program services, costs, and clients; and maintains records of quality reviews. When drawn together, this information could address many of the questions we have identified. The numerous evaluation studies concerning the Chapter 1 ESEA program similarly include information that covers many of our questions. Moreover, HRSA and the Department of Education's Office of Compensatory Education receive additional useful information—including clues to emerging concerns—through various program management activities as well as through active networking with service providers. In the sections below, we summarize what is available with respect to each of the information categories and questions listed in table 3.1.

Description

Both of these programs have gathered information from grantees and have used this information to provide aggregate descriptions and statistics, rather than to describe the range of variation to be found within the program. For Chapter 1 ESEA, data on clients, services, and student achievement are initially collected at the school or center level but are subsequently aggregated and reported at the state and national levels. CHC reports data only for the program as a whole. HRSA could piece together information from site visit reports and from its routine project management activities to describe for the Committee how community health centers and patients' experiences within them differ. Descriptive information of this kind is available from school case studies conducted in connection with the recent Chapter 1 ESEA evaluation.[4] The summary report of the evaluation included vignettes drawn from these studies and

[4]Sam Stringfield et al., Urban and Suburban/Rural Special Strategies for Educating Disadvantaged Students—First Year Report (Washington, D.C.: U.S. Department of Education, Office of the Under Secretary, 1994).

also presented data that was disaggregated to describe differences between high-poverty and low-poverty schools.[5]

Implementation

Information about how the CHC clinics implement recommended medical procedures is based on site visits, HRSA's quality review, and federal and regional officials' routine program monitoring. As mentioned previously, Chapter 1 ESEA has considered the quality of instructional practices to be a state and district responsibility. Federal program monitoring thus has not included the gathering of data on instructional quality. However, there are plans to make observation of instructional practices a larger part of federal monitoring in the future.

Targeting

The CHC program generates data about the numbers of clients served in its centers, and HRSA could develop population estimates to calculate the proportion of the targeted population that is served by each center and by the program as a whole in areas already designated as medically underserved. HRSA has relied on applicants to make known unmet needs in areas not currently included. (Community groups or local health organizations often apply for a designation of need.) In the past, it has not initiated its own assessment of medical need for areas from which there has been no application. However, based on Bureau of Primary Health Care (BPHC) estimates of areas with high priority for federal intervention, HRSA could determine the areas or people with the greatest general health care need and estimate the amount of unmet need in areas not now covered by CHCs. [6]

The 1993 Chapter 1 ESEA evaluation included a special study that examined the degree to which program funds were reaching the schools with the greatest concentrations of disadvantaged students, as well as the number of such schools that remain unserved. The question of whether funds are going to the most educationally needy students cannot be answered from grantee reports, but a study currently being conducted (the "Prospects" study) will shed light on this issue. (The final report from that study is due in January 1997.)

[5]Schools in which 75 percent or more of the students came from low-income families were designated high-poverty schools; low-poverty schools were those in which 0 to 19 percent of the students came from such families.

[6]HRSA is developing a proactive method of estimating the amount of need for primary care that does not restrict the focus to currently designated areas. Some recent data on the amount of unmet medically underserved needs are available from the National Association of Community Health Centers.

Impact

As noted previously, the primary purpose of the CHC program is to provide medically underserved people with access to medical services. Program data have focused on numbers served and service quality.[7] The express purpose of Chapter 1 ESEA, however, is to improve student achievement. Traditionally, two sources of information have been used to evaluate program impact. Results from the National Assessment of Educational Progress—which periodically tests a national sample of 4th, 8th, and 12th grade students in reading, writing, and mathematics—have been used to gauge whether disadvantaged students' achievement is improving overall. The program has also required states to report year-to-year gains in participating students' scores (aggregated to the state level) on nationally normed standardized tests.

These gain-score data give some idea of whether students in the aggregate are making progress, but they do not show the extent to which individual participants are benefiting from the program to the extent that they no longer need compensatory services. Nor do these aggregated data illuminate variations in impact across schools or subgroups of students. Attempts to use gain scores to identify schools that consistently fail to get results have been problematic, in part because of inconsistency in reporting and problems of data reliability.

In the future, the aforementioned "Prospects" study will fill in some of the gaps left by traditional information sources. That study will provide more differentiated information on services received and on achievement gains of Chapter 1 ESEA participants in comparison with other students. As of July 1, 1995, states will no longer be required to report nationally normed test score data. Rather, each state will assess student progress and adequacy of school performance in terms of its own standards for content coverage and student performance, using achievement tests and other instruments aligned to those standards. It is not yet clear what information will be reported to the federal program office.

Side Effects and Comparative Advantage

An issue for Chapter 1 ESEA has been whether providing compensatory education services through separate teachers or in separate settings has had undesirable side effects, such as impeding coordination between regular and compensatory instruction or encouraging classroom teachers to disclaim responsibility for Chapter 1 students' progress. While the program has data on how many schools use "pull-out" and other approaches and how much time students spend in each, these data do not

[7]HRSA collects some health outcome data and has recently contracted for an impact study of the CHC program. It has also supported efforts to develop health outcome measures pertinent to any primary care program.

illustrate how commonly used versions of each approach actually operate in relation to the regular classroom, how quality varies within each, and how variations in quality and approach are linked to student outcomes. The program has only begun to gather information at the level of detail needed to shed light on these questions and to compare the advantages of various models of delivering compensatory educational services. For the CHC program, information is available on program side effects. The comparative advantage of the CHC program over other programs has not been an evaluation issue, although it may be in the future.

What the Congress Requested and What It Received

The Congress did not mandate an evaluation study or reports on the CHC program. It did encourage evaluation of programs under the Public Health Service, which would include CHC, through legislation that set aside funds for that purpose. However, the set-aside provision did not require the agency to synthesize information and report it to the Congress. Reports have been provided when specifically requested, but this approach has left much useful information unreported, as the CHC case illustrates.[8]

The Committee informally requested studies and reports about particular aspects of the CHC program (such as the costs and effects of the Comprehensive Perinatal Care Program, quality of care, and capital improvement needs), and the agency provided them. In the absence of a request for a summary evaluation of the program, however, the bulk of the agency's information resources remained untapped. The only overall report the Committee received on the CHC program was the annual appropriations request—which stated the program's purpose, noted some of the services offered and the number of users, and reported previous program funding levels. However, this annual request, in keeping with the standard format for such requests (which are focused on appropriations issues), described the program in very broad terms. Committee staff commented that such general, highly aggregated information was not particularly useful for overseeing or evaluating the CHC program.

The Chapter 1 ESEA program offers a contrasting case in which the Committee has requested and received a great deal of information, both in interim years and for reauthorization. Soon after major changes were made to the program in 1988, congressional and agency officials became

[8]For information on the use of the PHS set-aside, see our earlier report entitled Public Health Service: Evaluation Set-Aside Has Not Realized Its Potential to Inform the Congress, GAO/PEMD-93-13 (Washington, D.C.: Apr. 8, 1993). The Preventive Health Amendments of 1993 (P.L. 103-183) amended the Public Health Service Act to establish a requirement for annual reporting of findings of evaluations conducted under the set-aside.

aware of implementation problems with the "maintenance of effort" provision for schoolwide programs.[9] The two sides communicated, and the problems were remedied. (We were told that this kind of interim-year problem identification and correction are common.)

The Congress mandated two major evaluation efforts for Chapter 1 ESEA: (1) a longitudinal study of a national sample of students (the "Prospects" study) and (2) an evaluation of the implementation and the effects of changes made in 1988. It also established a special panel to examine the use of standardized tests to assess student achievement.

In the course of its reauthorization deliberations in 1993 and 1994, the Committee received reports from the longitudinal study, the assessment study, and the numerous special studies that comprised the mandated evaluation. There was also a summary report that drew together findings and derived recommendations from all of these efforts. The plethora of study reports provided more information than the Committee could comfortably process. However, the summary report—which was organized around policy questions and provided easy-to-find summaries of main points, disaggregated data, and illustrative vignettes on case study schools—was viewed as useful.

The Congress also requested, and the Committee has received, a summary of Chapter 1 ESEA program participation and achievement data for each year—unfortunately, not until a year or more after the end of the school year for which data are collected.[10] General descriptive material and summary data are included in appropriation requests and in the Department of Education's annual summary of evaluations.

Communicating About Information Needs

In each of the cases we studied, we observed very limited communication between the parties concerning the Committee's questions and how the agency might best respond to them. Ad hoc inquiries aside, congressional requests for evaluative information took the form of statutory mandates. These mandates were quite specific, setting forth design details and report dates as well as objectives for the evaluation. The specifics were worked out in Committee deliberations with relatively little input from agency staff. In one of our cases, agency officials reviewed and suggested changes

[9]This provision concerns expenditures per pupil under a schoolwide program compared to expenditures in the same schools in the previous fiscal year.

[10]The agency has established new procedures and, as a result, expects to make data available more promptly in the future.

in an early draft that was under House committee review, but they did not discuss their concerns with committee representatives.

Once the evaluation mandates became law, agency officials in the Department of Education and HHS took them as a given. In preparing to implement the mandates, they went to considerable effort to interpret the interests and concerns expressed in the written record, but they did not confirm their interpretation with the Committee. Education Department officials did seek to obtain a sense of the relative priority of the questions listed and to discuss what they judged to be unworkable deadlines or design specifications, but Committee staff informed us that such behavior was the exception.[11] (In another case, the program manager did not see discussion of concerns with the Committee as a realistic option.) We found that, for the most part, agency officials simply did their best to do what was mandated.

Interchange was similarly limited as evaluations were conducted and reported. For example, Committee staff attended some of the meetings of the advisory group to the Chapter 1 ESEA evaluation, but as observers rather than as spokespersons for the Committee's information needs. Department of Education officials briefed the Committee on the forthcoming report, but Committee staff perceived that these briefings were designed to **inform** the Committee of what the Department had done rather than as an occasion for **dialogue** about priorities or about the best way to present the information.

Finally, we observed that the Committee rarely provided feedback to the agency concerning the reports it received. Committee staff did recall telling Department of Education officials that the clear organization and policy focus of the Chapter 1 report had been helpful. But they indicated that the Committee was unlikely to provide feedback on reports that were unsatisfactory—that is, that either failed to provide relevant information or buried it in an undigested mass of facts and observations.

The literature on public policy and the legislative process suggests that the communication patterns we observed are to be expected. The executive and legislative branches of government have different constituencies, institutional perspectives, and interests. Evaluative information often has policy implications that could have important consequences for the agency and department.

[11]The National Assessment of Chapter 1 was mandated in May 1990, with a final report due at the end of 1992. Meeting this schedule posed major challenges for the Department.

In view of their importance, policy-sensitive communications with the Congress are channeled through central congressional liaison offices, and contacts with committees are subject to high-level policy review and clearance. Discussions with our Committee respondents and agency officials, together with our review of agency policy, revealed that the liaison officer's chief role is to facilitate the exchange of information as legislation is drafted and enacted. Liaison officers also pass along ad hoc committee requests for information and arrange the briefings that precede the issuance of an evaluation study report. But they typically do not bring agency and committee staff together for evaluation planning.

Concluding Observations

Each of the three programs we studied collected a great deal of information relevant to questions of interest to the Committee. True, there were gaps in the information, and some useful questions could not be answered from available sources. However, we observed that lack of information does not appear to be the main problem. Rather, the problem seems to be that available information is not organized and communicated effectively. Much of the available information did not reach the Committee, or reached it in a form that was too highly aggregated to be useful or that was difficult to digest.

Second, we observed that the Committee was given no more than it asked for. It asked for no summary evaluation report on the CHC program, and it received none—either in interim years or at reauthorization. It mandated studies on a host of implementation and evaluation questions concerning the Chapter 1 ESEA program, and it received a report on every one. It mandated a report evaluating the CCDP, and it received such a report—although the report results could only be based on information from a few years of experience.

However, receiving what it initially asked for did not always meet the Congress' current needs, since these evolved over time (whereas statutory mandates remained fixed). For example, the Chapter 1 ESEA mandate called for a study of the implementation of services to children in private schools, but implementation was no longer a major issue by the time the report was due. The reauthorization and consideration of expanding Head Start services provided another example: it created a need for an earlier-than-scheduled report on emerging findings (on feasibility and implementation issues, in particular) from the CCDP program office.[12]

[12]Some of the emerging findings on implementation drawn from the report could have been fruitfully shared in a letter or briefing to the Committee while report editing and clearance proceeded as usual.

Finally, we observed that communication between the Committee and agency staff knowledgeable about program information was limited and comprised a series of one-way communications (from the Committee to the agency or the reverse) rather than joint discussion. This pattern of communication, which was reinforced by departmental arrangements for congressional liaison, affords little opportunity to build a shared understanding about the Committee's needs and how to meet them.

Taking into consideration what we found and reported in chapter 2, our case studies led us to conclude that obtaining timely and useful information for oversight and reauthorization requires not only **knowing** what questions to ask, but also **ensuring** that the information is in fact requested when needed and **communicating** with the agency to establish a mutual understanding of information needs and how they can be met. In chapter 4, we propose a strategy that incorporates each of these features.

A Strategy for Requesting Evaluation Information

Our final task for this report was to propose a strategy that the Committee might use to request information—one that would take into account the various factors that have contributed to unsatisfactory results in the past. The Committee's request letter recognized one such factor: the difficulty of knowing what questions to ask about programs of various types. Our strategy addresses that factor as well as two others that came to light through our case studies: agencies' disposition to provide information only when they have been asked to do so, and insufficient communication between the Committee and the agency with respect to information needs. We thus propose a strategy that attends not only to the content of information requests, but also to communication practices that can help ensure satisfactory responses to those requests.

The strategy we suggest for the Committee's consideration includes the following three components:

- selecting and adapting, from a core list, the descriptive and evaluative questions to be asked about a program in interim years and at reauthorization;
- arranging explicitly to obtain timely oversight information in interim years as well as to receive results of evaluation studies at reauthorization; and
- providing for increased communication with agency program and evaluation staff to help ensure that information needs are understood and that requests and reports are suitably framed and are adapted as needs evolve.

Identifying and Adapting Core Questions

We begin by proposing guidance to assist the Committee in identifying the questions to ask concerning a given program. In chapter 2, we outlined a set of core questions that might be asked about programs of various types. In the section that follows, we suggest how the Committee, drawing on its sense of the policy issues and of the history of a particular program, can select the core questions most important to that program, restate them in program-specific terms, and outline the forms of information needed to answer them. The Committee may elect to select and state the questions initially (as in the process of framing the questions for a mandated evaluation) and then request comments from the agency. Or, it may ask the agency to perform this initial step and to confirm the questions with the Committee.

In either case, we suggest that the Committee begin by reviewing the particular program's purpose and history, current policy issues regarding

this program, and the policy context. This initial review will help to determine which of the core questions listed in table 2.1 apply, what additional questions are of interest, and whether a full evaluation is needed. (It would probably be useful to ask the agencies to review the list also, as a first step in building a common framework for thinking about evaluation information.) For example, if policy review determines that the program's function could now be performed at the state level or that its dwindling population could be covered under a similar federal program, there might be little reason to seek additional agency information about program implementation or impact.

As noted in chapter 2, knowledge of a program's type or purpose is helpful in identifying the categories of information that are likely to be most critical to its evaluation. For example, the evaluation of demonstration programs (other than those that merely seek to provide an example) generally involves questions concerning implementation and impact. Statistics and research programs virtually always involve questions of implementation in accordance with professional standards and targeting to relevant problems.

Beyond this, programs must be reviewed on a case-by-case basis to determine which core questions are most pertinent. For example, the Committee or agency might ask itself the following questions:

- Do conditions within this program vary so substantially that the Committee will need disaggregated descriptive information?
- Are there new program aspects or provisions that are so significant that the Committee will want information on their implementation in interim years?
- Is targeting a critical issue for this program? If so, which of the targeting questions apply?
- Have possible side effects been identified that warrant a specific question?

In most cases, the Committee or the agency should restate each relevant core question so that it fits the program—for example, to identify the particular target population as "recently migrant students" rather than as "the appropriate people." The case studies reported in chapter 3 and in the appendixes provide examples of such restatements.

The Committee should also consider what form of information it needs in response to each question, and when and for what purpose (oversight or reauthorization) the information is needed. This is not to say that the

Committee should spell out in detail the specific data to be submitted. Rather, it should let the agency know its intent with regard to such questions as the following:

- Is the information needed for the next reauthorization or at an interim date before that time?
- Is an empirically precise answer—and hence time and resources to support a planned study—needed?

Examining its information needs in this way will help the Committee focus on the central issues, express its intent clearly, and avoid elaborate and expensive data collection when an estimate based on existing evidence would do as well. (For example, rather than developing an exact count of the number of sites that have implemented a new provision, program managers might estimate based on recent monitoring visits that about two thirds of the sites had done so.) This outline will form the basis for framing information requests that make the Committee's objectives clear, while avoiding detailed specifications that may prove to be unworkable. (Under our strategy, the agency is responsible for working out the details but must communicate with the Committee as it does so to ensure that the resulting plan meets the Committee's needs.)

To assist the Committee in creating such an outline, we have drawn on our observations from this study to create a guidance table (table 4.1). This table contains a row for each of our categories of information: description, implementation, targeting, impact, side effects, and comparative advantage. In addition, column one of the table lists our core questions in abbreviated form, column two notes the kinds of information that may be needed for each question, and column three notes the types of programs to which the question is likely to apply.

Table 4.1: Guide to Identifying Needed Information

Question	Information needed to respond	Applicability
Description		
Overall, what activities are conducted? By whom? How extensive and costly are the activities, and whom do they reach?	Narratives and summary data that describe the scope and extent of the program	All programs
If conditions, activities, and purposes are not uniform throughout the program, in what significant respects do they vary across program components, providers, or subgroups of clients?	Narratives and disaggregated data that show the extent of variation within the program	Programs that feature varied activities, clients, and settings
Implementation		
What progress has been made in implementing new provisions?[a]	Notification of early difficulties; descriptive data on the extent of implementation at specified intervals; answers concerning feasibility issues	All demonstration programs; other types of programs when new or when important new features are introduced
Have feasibility or management problems become evident?[a]	Same as above	Any program
If activities and products are expected to conform to professional standards or to program specifications, have they done so?	Answers concerning whether and to what extent criteria have been met; notification of evidence that standards or criteria may be inappropriate or outdated	All statistics and research programs; other types of programs when criteria are specified
Targeting		
Have program activities or products focused on appropriate issues or problems?	Description of the issues or problems targeted and evidence that supports their selection; if different components of the program target different problems, report them separately	Statistics, research, service, and regulatory programs
To what extent have they reached the appropriate people or organizations?	Answers showing how many of each relevant category of person or organization were reached; if targeting differs by site or by program component, report separately for each	Demonstration, service, and regulatory programs; statistics and research programs whose objectives include increasing the participation of or disseminating results to certain populations

(continued)

Question	Information needed to respond	Applicability
Do current targeting practices leave significant needs unmet (problems not addressed, clients not reached)?	Answers that show the nature and extent of critical unmet needs	Service or regulatory programs intended to cover a specified population fully; research and statistics programs intended to cover a problem area fully
Impact		
Overall, has the program led to improvements consistent with its purpose?	Answers concerning changes in outcomes and how they are related to program activities	All demonstrations; research, regulatory, and service programs that are intended to produce specified results
If impact has not been uniform, how has it varied across program components, approaches, providers, or client subgroups?	Answers concerning variations in conditions and program impact under each condition	Demonstration, service, and regulatory programs that deal with varied clients or varied conditions of implementation
Are there components or providers that consistently have failed to show an impact?	Answers concerning variation in impact across providers or components and extent of turnover in the low-performance category	Same as above
Side effects		
Have program activities had important positive or negative side effects, either for program participants or outside the program?[a]	Notification of possible effects; answers if evidence shows need for further investigation	Any program
Comparative advantage		
Is this program's strategy more effective in relation to its costs than others that serve the same purpose?.	Comparative analysis of effects of this program and of alternative approaches in relation to costs	Any program when alternatives and evidence concerning them are available

[a]Information important for interim years.

Arranging for Responses

Once questions have been identified, the Committee will need to arrange explicitly to obtain responses to them. Based on our observations, we suggest that the Committee set up procedures for reporting in interim years as well as arranging for program or comparative evaluations to be conducted in connection with reauthorization and policy review decisions.

Soliciting Interim-Year Information

Currently, there is no standard mechanism for prompting agencies to report information the Committee would find useful for oversight in interim years—that is, time-sensitive information that would keep the Committee informed of progress in implementing new provisions and of significant new evidence concerning program activities and effects. (As noted previously, neither budget submissions nor annual reports—as currently configured—are designed to serve this purpose.)

We propose that the Committee initiate a mechanism for prompting agencies (or a subset of agencies the Committee selects) to respond to program oversight questions each year. Agencies already prepare responses to questions about their budget requests each year, and they could provide responses to the authorizing committee on a similar schedule.

In the interest of evoking responses that are brief and up-to-date, we suggest that the Committee express its request for interim-year information through a letter (transmitted to the program agency through established departmental channels) rather than more formally through legislation.[1] The annual letter could include questions tailored to a particular program (such as progress in remedying a problem that was identified in a previous evaluation or oversight review). However, we believe that many of our core questions pertinent to interim years (for example, "what progress has been made in implementing significant new provisions" or "have feasibility problems become evident?") could be asked in general form in cases where the Committee's specific concerns have been communicated in earlier discussions with the agency.[2]

Agency officials indicated that they can respond to telephone or letter requests for time-sensitive factual information relatively quickly. However, responses to statutory reporting mandates or to Committee requests for information that raise budget or policy issues or that involve changes in legislation take time to develop. Such responses often involve policy or management concerns—concerns that go beyond the immediate program and that program staff are not authorized to address. Department officials appropriately require that these responses be centrally reviewed so that

[1]We expect that currently established departmental procedures would be followed in transmitting Committee requests to program agency staff and in approving and transmitting the latter's response.

[2]The questions to be included in the letter, and the timing of the requests, should be reviewed when the system of annual performance reports, established through the Government Performance and Results Act of 1993 (P.L. 103-62), is implemented in the year 2000. The fit between our approach and the act's is discussed in more detail later in this chapter.

these broader concerns can be addressed and input from different units coordinated.

It may be helpful if, along with stating the Committee's questions, the request letter or other communication included such language as: "Please contact (Committee staff person) if you have questions; we welcome the opportunity to discuss this request." This would encourage contact from the agency in cases where clarification was needed and would also give the department an opportunity to indicate which questions could be answered quickly and directly and which raised complex issues that would take more time to resolve.

Preparing for Reauthorization or Policy Review

As discussed in chapter 2, answering the evaluation questions that arise during reauthorization often requires data or narrative information beyond what the program routinely collects. We urge the Committee to notify agencies well in advance when it has such questions in mind so that they can plan (and budget) to obtain and analyze the necessary material in time. Such advance notice would help the agency set priorities for using the resources available to support evaluations. Comparative studies also require considerable advance preparation.

Communicating About Information Needs

The foregoing leads us to the final element of our strategy: increased committee-agency communication concerning information needs. As stated in chapter 3, present practice affords little opportunity for joint discussion and gives agency program and evaluation staff little basis for understanding what the Committee wants to know and what method of "packaging" the information would be most useful to it. We agree with other authors that increasing the opportunities for discussion would likely lead to more satisfactory results.

Increased communication is critical at two points and would be useful at several others. The critical points occur

- when the Committee frames a request for information, to ensure that the agency understands what is wanted and thus can alert the Committee to issues of content or feasibility that need resolution, and
- as report drafting begins, to assist the agency in understanding the issues that will be before the Committee and what kind of presentation format is thus likely to be most useful.

These points are critical because unverified agency assumptions or misinterpretations at either point can cause even the best-intentioned agency response to be off the mark.

The additional points at which we suggest communication occur

- when the agency develops the design of a major study, to verify that it is on target from the Committee's perspective;
- midway through the period between reauthorizations, to consider whether issues and priorities have changed and to involve any new Committee staff; and
- after a report is submitted, to convey feedback concerning its usefulness to the Committee.

Communication at these points might not be necessary in cases where the task before the agency is clear-cut, the issues are unlikely to change, and the report is well received. However, in less favorable circumstances, these communications points would likely be useful.

Additional communication can be facilitated by using organizational structures and procedures that are already in place. It is important for the Committee to initiate requests and to give agencies a clear signal that it welcomes discussion with the agency, as previously outlined. The congressional liaison officer for the agency can then assist by providing agency program and evaluation staff with background information about the Committee's intent. He or she can also bring agency and Committee representatives together so that issues can be resolved and mutual understanding achieved.

Feasibility of Our Approach

During the course of our study, we discussed our proposed strategy with Committee staff and with officials in the agencies that served as case studies. Both groups found our core questions useful and acknowledged the potential benefits and even a common interest in moving in the direction we propose.

Both groups also noted factors (ranging from demands on Committee time to potential conflicts over where to draw the line between legislative oversight and executive responsibility) that create obstacles to the kinds of collaboration the strategy envisions. Given the different constituencies, institutional perspectives, and interests of the Congress and the executive

departments, these obstacles are to be expected. However, we were impressed by the commonality of interest that we observed between the two groups. We believe that it constitutes a solid basis for moving forward to improve the usefulness of evaluation information.

Agency Comments and Our Response

In written comments on a draft of this report, the Departments of Education and HHS agreed to the benefits of increased consultation between the Committee and the agency regarding evaluative information needs. However, they were concerned that our proposed strategy be consistent with the requirements of the Government Performance and Results Act and not constitute an added burden. The Department of Education also expressed concern about two of our categories of information—side effects and comparative advantage. The Department noted that questions under these categories can be difficult for the program agency to address if they require information about programs administered by another agency. We address these concerns in turn.

Our Approach and GPRA

While this study was in progress, the Congress passed the Government Performance and Results Act of 1993 (P.L. 103-62). By the turn of the century, annual reporting under this act is expected to fill some of the information gaps we describe in this report. Among other things, GPRA requires every agency to establish indicators of performance, set annual performance goals, and report on actual performance in comparison with these goals each March beginning in the year 2000. Agencies are now developing strategic plans (to be submitted September 30, 1997) articulating the agency's mission, goals, and objectives preparatory to meeting these reporting requirements.

We compared our approach with GPRA's and concluded that the two are compatible. Both approaches emphasize the importance of agency consultation with the Congress as evaluation strategies are planned, goals and objectives identified, and indicators selected. Both note the importance of providing information that indicates how well the program is doing with respect to its intended objectives, and both call for annual reporting of information to provide the basis for accountability and effective oversight.

We foresee little added burden—and some efficiency of effort—from using our approach in conjunction with GPRA. The procedures for identifying and adapting core questions presented earlier in this chapter are likely to be

useful in preparing to meet GPRA requirements, and consultation with the Committee should ensure that data collected to support GPRA reporting can also be used to meet the Committee's special needs (for example, that performance data can be disaggregated in ways that are important to the Committee).

At the same time, our suggested approach provides a useful complement to GPRA. Our approach to requesting interim-year information can be applied immediately, to take advantage of currently available information while GPRA reports are being planned. Moreover, while GPRA annual reporting focuses on intended program outcomes, our approach covers additional categories of information—description, side effects, and comparative advantage—that the Committee finds useful, especially in connection with major program reauthorizations or policy reviews.

Evaluating Side Effects and Comparative Advantage

We agree that to evaluate comparative advantage typically requires data beyond what the administering agency can provide and that some side-effects questions might also pose this problem. Our discussion in chapter 2 notes that by virtue of their complex informational requirements, comparative advantage studies typically are implemented through special arrangements. Special arrangements would similarly be appropriate for conducting a detailed evaluation of one program's side effects on other programs.

Many side-effects questions, however, fall within the scope of evaluation for a particular program (as our case study examples illustrate). The process of discussion that we have suggested should help the Committee and the agency identify the side-effects questions that are pertinent to a program and the level of detail or precision needed for each. Discussion should also examine whether it is feasible for the administering agency to gather this information or whether special arrangements will be needed.

Comprehensive Child Development Program

Program Description

Program Authorization and Objectives

The Comprehensive Child Development Program was authorized in 1988 under the Comprehensive Child Development Centers Act of 1988 (P.L. 100-297). CCDP was reauthorized through fiscal year 1994 under the Augustus F. Hawkins Human Services Reauthorization Act of 1990. In May 1994, elements of CCDP were incorporated into a new Head Start Program for Families With Infants and Toddlers, authorized under title I of P.L. 103-252, the Head Start Act amendments.[1]

CCDP was designed to enhance the physical, emotional, and intellectual development of low-income children (infants to school age) and to contribute to self-sufficiency by providing support to their parents and other family members. This demonstration program's purpose is to test whether it is feasible and effective to provide integrated and comprehensive support services early in the child's life and within the family context.

Program Operation, Participation Levels, and Funding

The Department of Health and Human Services funded 24 CCDP grants in 1989 and 1990 and 10 more in 1992. Each grant provided funding for 5 years. There are currently 34 projects in rural and urban areas providing core services (health care, mental health care, child care, early educational intervention, early childhood development, prenatal care, parenting education, employment counseling, vocational training, adult education, and nutritional assistance).

A CCDP grantee organization (which may be a health clinic, a family services agency, a university, or even a school district) acts as a service integrator by building supporting networks with community agencies and at times facilitating and advocating for services. Each grantee organization assesses local needs in its area and plans ways to meet these needs through direct intervention of family-focused case management and a combination of project-provided services and services arranged through third-party providers. All projects use the case management approach; however, the degree of reliance on existing agencies in the community

[1]This appendix describes CCDP until the May 18, 1994, Head Start Act amendments and reauthorization.

varies from project to project. Each project has an advisory board with community, business, and client-family representation.[2]

Fiscal year 1991 funding for this program totaled $24 million; for fiscal year 1993, an increase of $23 million ($20 million received from the Department of Education) brought the total to $47 million, which was also the amount requested for fiscal year 1994.

Administrative Structure

The Head Start Bureau of the Administration on Children, Youth, and Families in the Department of Health and Human Services is responsible for the administration of CCDP grants. ACYF's functions include selecting grantees, providing technical assistance, monitoring their compliance with program regulations, and directing contractors in the performance of process and impact evaluation studies.

Mandated Evaluation

The authorizing legislation for CCDP mandated an evaluation report on program impact and program feasibility that was due on October 1, 1993. HHS provided an interim evaluation report—on short-term program impact and on the structure and mechanisms of service delivery—to the Senate Committee on Labor and Human Resources in May 1994. That interim report covered the results of approximately the first half of the 5-year cycle of services. A final report covering the entire program is scheduled for delivery in March 1996.

Major Sources of Information

To meet the evaluation requirements, the agency has arranged to collect diverse types and sources of information. The grantee organizations are required to collect data on program operations, user participation, and compliance with program regulations. In addition, each project submits progress reports on its activities (including vignettes on participants' successes) to ACYF. A nonparticipant observes agency operations and records information on side effects and local community dynamics in ethnographer reports. ACYF conducts site visits that result in a letter assessing regulatory compliance and implementation and discussing quality issues (site visit reports). These program-monitoring documents also contribute to the assessment of CCDP operations and feasibility.

[2]ACYF refers to individual CCDP grantee sites as "centers." We use the term "project" instead to refer to the physical center itself, community characteristics, and services available in the community.

In December 1991, the agency published a report that described CCDP and start-up and enrollment issues (referred to as the CCDP First Annual Report). The agency has recently published its interim evaluation report tc the Congress entitled "Comprehensive Child Development Program—A National Family Support Demonstration" (referred to as the CCDP Interim Report). The CCDP Interim Report provided analyses of program implementation and short-term program impacts on families and communities, as well as a description of families and projects. In a final report to be prepared after the enrolled families have received 5 years of services, the agency expects to provide a greater understanding of how effects vary with the intensity and duration of services, as well as with differences in the effects of projects' characteristics. It will also provide a picture of the effects of long-term participation on families and children.

Evaluation Questions and Information Needs

Interim-Year Questions

CCDP projects have reported on start-up problems and early progress. The Committee might also ask to be informed of new feasibility issues and of program impacts or significant side effects (discussed in more detail below) as they become evident at existing and new projects under the reauthorized Head Start Program.

Reauthorization Questions

After the Head Start reauthorization of 1994, CCDP will not continue as a separate program, but its grantees will continue to operate, together with new ones, under the newly authorized Head Start Program for Families With Infants and Toddlers. Because this new program appears generally similar to CCDP, we expect that the questions posed for CCDP could also inform oversight of the new and continuing projects.

Table I.1 suggests evaluation questions pertinent to a demonstration program that are drawn from the list of core questions in chapter 2 and restated in terms of CCDP. The table also identifies indicators or measures needed to answer those questions and summarizes related information that is currently available. It demonstrates the use of the question-selection strategy proposed in chapter 4. The questions and indicators are suggestive and not prescriptive for the program. Under our

strategy, the Committee and Department staffs would collaborate to work out the specific questions and indicators needed.

Table I.1: CCDP Evaluation Questions and Related Information

Question	Indicator or measure	Available information
Description		
What model or approach is tested here (core services provided, service delivery method, organization)?	Description of required elements of the model	Requirements are stated in authorizing legislation and program documents and were summarized in the CCDP Interim and First Annual Reports.
How do features of the model differ across projects?	Narrative descriptions of major variants and data showing the frequency with which each occurs	Management information system, progress reports, and ethnographer reports provide relevant information; CCDP Interim Report notes significant variants on the model (for example, type of grantee agency, staffing configurations, and interagency agreements).
What are the conditions (project size and philosophy, client makeup, or community resources) under which services are provided, and how do they differ?	Range and amount of variation in project characteristics that affect implementation of program model	Management information system, progress reports, and ethnographer reports provide this information, but the CCDP Interim Report does not compare critical project-based characteristics of program operations; the final report is expected to make such comparisons.
Implementation		
What start-up and operating difficulties did projects face and how were they resolved?	Observational or participant survey data collected as program was implemented	Site visit, ethnographer, and progress reports described difficulties of putting the new projects into operation and providing all core services; both the CCDP First Annual and Interim Reports cover start-up and operational difficulties and early solutions.

(continued)

Question	Indicator or measure	Available information
What modifications in the model were made by the grantees or program (for example, in case management or in arrangements with other service providers)?	Observation or survey responses describing the modifications and reasons for making them, descriptions and classification of arrangements between organizations at different times	Progress, site visit, and ethnographer reports, as well as refunding grant applications, contain pertinent information, some of which is summarized in the CCDP First Annual Report; the basic program model is also described there.
What has CCDP demonstrated about the feasibility of applying the program model?	Data that illuminate conditions for successful implementation, barriers to implementation, and cost of implementation	The sources mentioned above, supplemented by grantees' financial records, include the needed information; CCDP Interim Report provides information on implementation problems and some solutions in the early years of the program.
Targeting		
Question is not applicable—client selection not discretionary	Not applicable	Not applicable
Impact		
To what extent did services improve child development and family self-sufficiency?	Measures of children's and family members' skills and behaviors upon enrollment and after receiving services; measures of services received; comparable data from comparable families that did not participate.	CCDP Interim Report provides information on a variety of outcomes for the child, mother, and family as well as for members of the control group; these data can be combined with descriptive data about users and services to provide answers about impact.
Under what conditions, and with what kinds of services or delivery mechanisms, has the program been most or least successful?	Measures of child and family outcomes by project and service, displayed to show range and variance.	CCDP Interim Report provides programwide child and family outcomes; subsequent reports are expected to describe how effects vary by services provided and project-based characteristics.

(continued)

Question	Indicator or measure	Available information
Side effects		
Have CCDP activities prompted increased provision, quality, and coordination of services to the non-CCDP community?	Evidence of addition of new services or increased coordination among existing services as an outgrowth of project networking efforts	Ethnographer, advisory board, and project director's reports captured these observations, and positive changes were noted in CCDP Interim Report.
Comparative advantage		
Would a reduced version of CCDP (which included only the most viable elements) be more effective than other approaches?	Estimates of enhanced family self-sufficiency and child development under modified-CCDP, impact data from comparison program or other approach	CCDP evaluation studies will provide some impact data, which would be basis for estimates of modified-CCDP effects; a separate study is needed to compare these estimated effects with the effects of other programs.

Community Health Centers

Program Description

Program Authorization and Objectives

The Community Health Center (CHC) program was authorized in 1975 under section 330 of the Public Health Service Act. At the time this report was written, the program was authorized through fiscal year 1994, and reauthorization for fiscal year 1995 was expected shortly. The program's purpose is to provide access to comprehensive and case-managed primary health care services to rural and urban populations living in medically underserved areas.[1] The program also targets "at-risk" populations (pregnant women, children, substance abusers, and elderly persons), and centers seek to overcome barriers to health care access, such as those related to cultural and language differences.

Program Operation, Funding, and Participation Levels

The CHC program funded 579 grantee centers serving 1,575 sites on a budget of $558 million in fiscal year 1993. (The fiscal year 1994 appropriation was about $603.5 million.) All centers provide health care services such as physician services, diagnostic laboratory and radiological services, and pharmaceutical and emergency services, as well as preventive care like immunizations, dentistry, family planning, and vision and hearing screening. Other required services include translation services, transportation, and referrals to other providers, and centers can opt to provide supplemental services like health education and outreach. Services are provided to about 6 million medically underserved or "at-risk' people a year. The fiscal year 1994 appropriation of $604 million represents an increase of $46 million to fund new sites and serve more people.

CHC centers serve areas (urban and rural) with poor access to medical services, and they also provide supporting services targeted to the local community's needs (such as transportation, day care, or culturally sensitive care to minority groups). Most centers operate on a fee-for-service basis, while others provide managed care.

Administrative Structure

The CHC program is administered by the Bureau of Primary Health Care. BPHC is part of the Health Resources and Services Administration within

[1]An area is designated "medically underserved" by HHS, in conjunction with state authorities, based or a formula that includes four variables: physician-to-population ratio, infant mortality rate, poverty population, and population 65 years of age and older. An area may encompass a neighborhood or an entire county.

the Public Health Service of HHS. Regional HHS offices assist headquarters in monitoring grantees' compliance with program regulations. The CHC has a subprogram, the Comprehensive Perinatal Care Program, that provides additional funding for enhanced perinatal services and is treated here as a separate program.

Mandated Evaluation

The last reauthorization did not mandate an evaluation or any type of report to the Congress on this program. However, HRSA has initiated and conducted several evaluation studies, which are discussed in the next section.

Major Sources of Information

The legislation requires (1) that community health centers have an effective procedure for collecting information on their costs of operation, (2) that information on the use of services (their availability and accessibility) be collected, and (3) that grantees collect data on program operations and user participation. HHS regional offices collect and enter computerized data received from grantees into the Bureau's Common Reporting Requirements (BCRR) data system. BPHC is planning changes in this system to allow collection of additional data on clients, their health status and needs, and types of services provided. A new provision is intended to result in information about the age-appropriate preventive services provided at different sites. Grant applications include information on each center, the health needs of its community, and progress toward achieving its program goals (such as retaining clients).

BPHC, with the aid of HHS regional offices and on-site assessments conducted by federal staff and consultants, reviews center operations and compliance with BPHC-mandated budgetary and performance standards (such as following standard medical procedures and appropriate staff licensing guidelines). This is intended to ensure that centers are properly managed and operated. The bureau also has information on special populations' needs and nationwide health needs derived from individual grant applications and a variety of other sources such as the U.S. census, national health surveys, and state and local organizations.

BPHC has conducted several evaluation studies and is planning others. The completed studies have provided baseline information on the range of preventive services offered, examined the effects of Federally Qualified Health Center provisions on individual centers and the capacity of each to expand its service, and traced the impact of CHC use on some Medicaid

costs and hospitalizations. Ongoing studies will provide assessments of the case management approach and the managed care system used in certain centers. Several planned studies should provide in-depth descriptive profiles of CHC users and of services provided, as well as an evaluation of program effectiveness and whether CHCs do improve the health status of their users.[2]

Evaluation Questions and Information Needs

Interim-Year Questions

The forthcoming reauthorization may introduce new provisions that raise implementation questions for the next few interim years. Beyond these, questions concerning the emergence of significant feasibility or service quality problems and new evidence concerning program targeting, impact, or side effects will be relevant.

Reauthorization Questions

The CHC program is now scheduled for reauthorization consideration in 1995. Table II.1 suggests evaluation questions pertinent to the CHC program that are drawn from the list of core questions in chapter 2 and restated in terms of this program. The table also identifies indicators or measures needed to answer those questions and summarizes related information that is currently available. This table demonstrates the use of the question-selection strategy proposed in chapter 4.

[2]These studies will also evaluate the Migrant Health Center program and, in one instance, will combine the two programs and their results. Since the CHC program funds some centers that also get Migrant Health program funding, the programs are usually located and operated together. Roughly 100 jointly funded centers provide the same basic services to a mix of migrant and resident clients using the same staff, thereby making it difficult to identify unique CHC program operations and effects. About 450 organizations are CHC only, and 20 are Migrant Health only.

Table II.1: CHC Evaluation Questions and Related Information

Question	Indicator or measure	Available information
Description		
What services are actually provided, whom do they reach, and in what kinds of settings?	User profile of client characteristics (insurance status, relative income, ethnicity, age, sex); profile of services provided (required services and range of optional services)	User profile information is accessible in BCRR; information on required and optional services is provided; and information on center settings is available in grant applications and BCRR.
How do centers differ in terms of location (rural or urban), client characteristics, and area population?	Profile of a sample of centers, noting variation in each feature across centers	Center location and description are available in grant applications, and client demographics are accessible in BCRR.
What other aspects of the program are significant for understanding program management and use of federal money?	Information on relevant variables (such as payment systems used, percentage of clients covered by insurance, and delivery models)	Information on these features is available from BCRR and grant applications.
Implementation		
How have recent increases in funding affected program implementation with regard to capacity to serve, services provided, and location?	Number of new sites or expansions; number and description of newly served clients	Information is accessible in BCRR and program files.
Does evidence suggest significant compliance problems with centers' mandate to provide minimally adequate care or with regard to any statutory requirement?	Frequency and severity of reports on significant center problems or noncompliance with critical requirements	Information on important grantee noncompliance may be collected from diverse sources during routine program monitoring.
Are there serious problems that centers face in implementing the program?	Program management information suggesting problems, noting frequency and severity of situations such as unmet staffing needs, facilities needing renovation, or high insurance costs	Information about the problems centers face is available in grant applications and from the annual center review; headquarters or regional staff may also be aware of other relevant information.
Targeting		
To what extent are centers covering their medically underserved populations or areas?	Portion of population in medically underserved areas or counties that CHCs serve	Medically underserved population size is available from grant application and number of users in BCRR, so portion can be calculated.

(continued)

Question	Indicator or measure	Available information
Are centers distributed around the country to cover the most seriously underserved populations or areas?	Proportion of medically underserved areas in U.S. without a center; identification of those areas without centers that cover the most severely underserved counties (among all counties)	Amount of medical need in a specific area is available from grant applications. Amount of need in unserved areas is available from various national and local sources; agency has reported amount of need met by CHC program but not its size relative to the amount of unmet national need or the severity of need; NACHC has reported on the amount and severity of need.
To what extent do centers serve special population clients?	Proportion of clients who are special population members	Number of users, including some special population users (the elderly, children, infants) is available from BCRR; information on each center's assessment of services needed and services provided is available in grant application and BCRR.
Are there major client needs that center services do not address?	Evidence that centers provide services that match needs identified in the grantee's needs assessment; evidence that program services offered match program services needed	Information on each center's assessment of services needed and services provided is available in grant application and BCRR.
Impact		
Have CHC centers with optional, culturally sensitive, and other outreach services succeeded in increasing access to health care for the targeted populations?	Change in coverage of clients and increased utilization by clients; change in proportion of ethnic group and other-language members served	BCRR has relevant client information over time; evaluation or public health literature may report on program outreach effects.
Has the medically underserved population served by a center achieved higher health status?	Information on continuity of client care, reduced use of emergency room facilities for nonemergency care, and change in morbidity in area	Health status information, use of health care facilities, CHC and non-CHC client information are available; agency has ongoing and completed evaluation studies regarding these issues.

(continued)

Question	Indicator or measure	Available information
Side effects		
Have there been effects on other health care providers?	Changes in use of emergency facilities, hospitals, or school-based care or nursing	Grant application or ad hoc sources may report this information; evaluation study on change in Medicaid usage is available.
Comparative advantage		
Is providing clinics (as opposed to encouraging others to provide service) more successful in increasing access to care for the targeted populations?	Comparison of costs, types of services provided, and clients served between CHCs and another nonclinic delivery approach	Agency has similar evaluation studies planned.

Chapter 1 Elementary and Secondary Education Act

Program Description

Program Authorization and Objectives

The Chapter 1 program (originally known as Title I) was established by the Elementary and Secondary Education Act of 1965. This summary describes the program as it existed under the 1988 Augustus F. Hawkins-Robert T. Stafford Elementary and Secondary School Improvement Amendments (P.L. 100-297). (Substantial changes were introduced in 1994, as discussed later in this appendix.) Chapter 1 provides financial assistance to local education agencies to help educationally deprived children in low-income areas (1) succeed in the regular program of the local education agency, (2) attain grade-level proficiency, and (3) improve achievement in basic and more advanced skills.

Program Operation, Funding, and Participation Levels

The program primarily provides compensatory instruction in basic subjects (reading, mathematics, language arts). Some supporting services (social work, health and nutrition, transportation) are also provided. In most schools, only low-achieving students receive assistance, which may be given in the classroom or via "pull out" to a separate setting. Schools in which 75 percent or more of the students come from low-income families may use Chapter 1 funds to support schoolwide improvement rather than serve only the low achievers. Eligible students who attend private schools receive services provided off-site through the public schools. Parent participation is encouraged.

With an annual budget of $6.7 billion, Chapter 1 is the largest federal program in elementary and secondary education. Grants are awarded to state education agencies, and through them to school districts in accordance with the numbers of students from low-income families. All states, Puerto Rico, the District of Columbia, and the outlying territories participate. Funds flow to 90 percent of all school districts, over 70 percent of all public elementary schools, and 30 percent of public secondary schools. Nearly 5.5 million children are served.

Administrative Structure

Chapter 1 is administered by the Office of Compensatory Education Programs within the Department of Education's Office of Elementary and Secondary Education. Federal program staff review the work of state

Chapter 1 program directors, who in turn review compliance and progress at the local level.

Mandated Evaluations

The Chapter 1 program was the subject of major mandated evaluations in 1977, in 1987, and again in the 1992 National Assessment of Chapter 1. The 1992 assessment drew on preliminary findings from the "Prospects" study, a national longitudinal study mandated in 1988 to trace the effects of Chapter 1 participation to young adulthood. The Congress has also requested GAO reports on various aspects of the program.

The 1988 legislation required states to evaluate their Chapter 1 programs at least every 2 years. State review of student gains for each school in terms of standardized test scores was deemed sufficient to meet this requirement and was used to identify schools in need of improvement. Local education agencies were to conduct program evaluations every 3 years. At the school level, schoolwide projects were to be evaluated after 3 years of operation. In addition, individual students' gains were to be examined and individualized educational plans devised for those who were not making progress.

Major Sources of Information

The 1992 National Assessment drew on numerous studies of special issues and populations as well as case studies of school programs. These are summarized in its final report.[1] It also drew on preliminary results from two major studies: (1) the "Prospects" study, which examines the school experiences and achievement of nationally representative samples of disadvantaged 1st, 3rd, and 7th graders (some of them receiving Chapter 1 services, and others not) over many years, and (2) an observational study entitled "Special Strategies for Educating Disadvantaged Students."

Annual state education agency reports to date have listed the number of schools, staff, and students participating, as well as the number of schools deemed in need of improvement. State agencies also reported the statewide average pretest and posttest scores for Chapter 1 students.

As its basis for allocating funds to states, the program has relied primarily on family income data from the decennial census. Local education agencies have used data on numbers of poor children from the Department of Agriculture's school lunch program to determine school

[1]U.S. Department of Education, Reinventing Chapter 1: The Current Chapter 1 Program and New Directions (Washington, D.C.: Feb. 1993).

eligibility. Although it cannot distinguish participants in Chapter 1 from other students, the National Assessment of Educational Progress provides general evidence of how disadvantaged students' achievement in basic academic subjects compares with that of more advantaged students.

Changes Enacted in 1994

The Improving America's Schools Act of 1994 (P.L. 103-382) renamed the Chapter 1 program as Title I and introduced significant changes. The 1994 law specifies components or qualities that funded activities should incorporate—including high standards, enriched programs, upgraded instruction, and improved teacher professional development—and alters the kinds of information that will be available. Under this law, the requirement that student achievement be measured in terms of performance on nationally normed standardized tests has been dropped. Instead, consistent with the provisions of the Goals 2000: Educate America Act (P.L. 103-227), states are encouraged to evaluate student achievement and school performance in terms of standards and assessments specific to each state and applicable to all students, not just those participating in Title I programs.

In place of the historic emphasis on providing compensatory services to individual students, the Improving America's Schools Act emphasizes strengthening the regular program in schools that serve low-income students. It would permit many more schools to adopt schoolwide programs (in which funding is used to upgrade the entire educational program and "Title I students" are not distinct from other students). Student achievement and school improvement would be evaluated only at certain grades selected by each state (such as 3, 8, and 11). Thus, student achievement data will be cross-sectional and state-specific. Data will be drawn from a range of students who attend schoolwide programs and from low achievers who have received special services in schools with "targeted" programs. At the time of this writing, future reporting requirements had not yet been worked out.

The act provides for data on individual students' year-to-year progress to be collected through a national longitudinal study. The longitudinal study will compare student performance against content standards of national scope. A second mandated evaluation study will examine progress toward the goal of having all children served by this program reach their state's content and performance standards. This evaluation will also examine the targeting of resources and the extent to which the high standards, enriched programs, upgraded instruction, parental participation, improved

teacher professional development, and other program practices specified in the act have been put in place. Projects to demonstrate effective practices are also authorized.

Evaluation Questions and Information Needs

In identifying the Committee's future information needs, we took into account the major changes adopted in 1994. Specifically, we assumed (1) that school and student performance would be evaluated in terms of standards and assessments specific to each state rather than in terms of nationally normed tests, (2) that only cross-sectional data for certain grades would be available from state records, and (3) that longitudinal data on a sample of individual students would be available from the "Prospects" study and from the newly authorized longitudinal study. Information may also be available from projects funded to demonstrate effective practices.

Interim-Year Questions

In light of these changes, in the years immediately following the 1994 reauthorization, the Committee will likely wish to be kept informed of

- progress in implementing the new requirement for state standards and assessments and related changes in curriculum, instruction, and staff development;
- feasibility issues that have arisen as implementation has moved forward;
- early evidence that new assessments meet criteria of technical soundness and are aligned to state standards;
- early evidence of effects of new standards and assessment practices on student achievement; and
- early evidence that the changes have unanticipated side effects on instruction or on parents, teachers, or students (especially students with disabilities or limited English proficiency).

Reauthorization Questions

Table III.1 lists evaluation questions that the Committee may wish to ask at the next ESEA reauthorization (scheduled to take place in approximately 5 years), indicators or measures needed to answer them, and related information that is currently available. The questions are drawn from the list of core questions in chapter 2 and restated in terms of this program. As noted earlier, it is unclear what information about student and school performance will be reported to the administering agency under the new provisions. The question of whether and how state data can be aggregated to form a national picture, given that each state uses a different standard,

has yet to be answered. The entries under "Impact" in table III.1 suggest one possible method of accomplishing this task. We consider our approach illustrative of what **might** be done. Under our strategy, Committee and Department staff would collaborate to work out an approach.[2]

Table III.1: Chapter 1 ESEA Evaluation Questions and Related Information

Question	Indicator or measure	Available information
Description		
How different are participating schools and districts from one another in terms of socioeconomic variables, student performance, and school resources?	High, low, and modal value(s) for each feature; typical (frequently occurring) combinations of features	Socioeconomic data available from free or subsidized lunch program; achievement data from "Prospects" study; resource data from special studies
How do schools differ in instructional strategy and method of delivery and in the added services provided to low-achieving students?	Narrative description of major variants and the school conditions under which they occur; frequency count or estimate of relative frequency of each variant	Capsule descriptions of successful schools, case studies of school practices, and observations drawn from site reviews; "Prospects" study may contain data
Implementation		
How many states have established content and performance standards, and how widely do these standards vary?	Number of states that have standards; examples or illustrations of standards that differ in scope and level	Standards to be set forth in each state's plan; additional details may be needed for adequate description
To what extent are state assessments aligned to state content and performance standards, of acceptable technical quality, and appropriately adapted for limited English proficiency students and students with disabilities?	Findings from reviews of instrument content, administration, and scoring methods; evidence of reliability and validity of results	Federal and state quality criteria and review procedures not yet established

(continued)

[2]Consistent with the 1988 as well as the 1994 law, this program's central goal is to bring participating students up to a certain level of educational performance. Thus, its impact is properly evaluated in terms of how many students, and which kinds of participating students, achieve this goal. The program also is intended to lead states, school districts, and schools to adopt improved practices. We treat these practices as standards for program operation, an aspect of implementation.

Question	Indicator or measure	Available information
What proportion of participating schools have programs that exhibit the characteristics of quality education embodied in the authorizing statute? What proportion lack the essential characteristics of this model?	Frequency data from school surveys and from observations of actual school practice	Future evaluation studies authorized to collect relevant data, but such studies not yet designed; state or other site reviews also a potential source
Targeting		
To what extent are services targeted to schools with high concentrations of low-income students?	Proportion of participating schools in which there are high concentrations of students from low-income families; proportions of participating students in low, moderate, and high poverty schools	Data available at the school level but may not be reported; "Prospects" study includes relevant data
To what extent are services reaching low-achieving students in these schools, as well as students with special educational needs (such as limited Enghlish proficiency, disability, or high mobility)?	Proportion of students served/benefited who are low achieving, have special needs, or both	Data on achievement, handicap, and English language proficiency for each student served available at the school level, but not yet clear whether these data to be reported; "Prospects" study contains relevant data
How many severely needy schools and students remain unserved by this or other comparable program? Have numbers and percentages decreased since 1994?	Number of unserved schools, by percentage of low-income and low-achieving or multiple-needs students; number of unserved needy students in all eligible schools	Local education agencies gather school poverty data for school selection but need not report it; National Center for Education Statistics Schools and Staffing Survey contains relevant information
Impact		
In the aggregate, is the performance of economically disadvantaged students approaching that of their advantaged peers?	Measure of student achievement and economic status for national sample, such that distributions can be compared	National Assessment of Educational Progress provides this kind of information

(continued)

Question	Indicator or measure	Available information
Are students served by this program being brought to the level of performance expected of all students?	Percentage of students whose performance (1) is seriously deficient and shows little improvement, (2) has improved but is still below expected level, (3) has reached the expected level but student needs assistance to maintain gains, or (4) is sufficiently strong to "graduate" from Chapter 1	Schools classify students' progress and achievement in order to determine who needs extra assistance; however, these judgments are not reported, and it may be some years before they can be linked to state standards; new longitudinal study may gather relevant data
	Percentage of students who have received Chapter 1 services for 3 years or more and have not reached the expected level	"Prospects" study to have length-of-participation data for a sample of students
Under what conditions, and with what kinds of students, has the program been most or least successful? Are schoolwide programs more effective for low achievers than programs that target specific students?	Student performance data linked to school, student, and service delivery data	"Prospects" study to provide information; additional data will be needed
Has the school improvement procedure led to improved student outcomes?	Number of schools needing improvement and percentage in which student performance has improved substantially; comparative data on performance of low-achieving students in schools that serve similar populations but differ in type of program	Criteria for school improvement to be set by states but are not yet established
Side effects		
Has the adoption of schoolwide programs decreased the extent or effectiveness of assistance provided to low-achieving students in those schools?	Observational data on extent of services to low-achieving students prior to and after adoption of schoolwide programs; if not available, cross-sectional comparative data	"Prospects" study may provide some data but may include too few cases
Has the nature, amount, and cost of student testing changed substantially?	Percent of districts that drop annual testing, cease to use a nationally normed test, or use either performance-based tests or multiple assessments; per-student cost of testing	New issue; what information will be available not yet known

(continued)

Question	Indicator or measure	Available information
Where the nature of assessment has changed, to what extent have parallel changes occurred in instruction?	Proportion of classrooms in which such changes have been reported or observed	New issue; what information will be available not yet known
Has use of standards-based testing adversely affected students who face special barriers (language, handicap) to meeting such standards?	Changes in proportion of such students who are included in assessment; patterns of attendance, effort, and persistence in school on the part of such students	"Prospects" study may provide some data but additional information will likely be needed
Comparative advantage		
Is there evidence that tying Chapter 1 criteria to state standards has been more advantageous in relation to its cost than the prior approach? (Generally, or only in some states?)	Change in rate of aggregate achievement of disadvantaged students nationally and in each state; changes in costs; case study data from varied states	National and state results from the National Assessment of Educational Progress may be useful; plans for other data collection not yet known

Comments From the Department of Education

UNITED STATES DEPARTMENT OF EDUCATION

OFFICE OF ELEMENTARY AND SECONDARY EDUCATION

THE ASSISTANT SECRETARY

October 19, 1994

Ms. Terry E. Hedrick
Assistant Comptroller General
Program Evaluation and Methodology Division
United States General Accounting Office
Washington, D.C. 20548

Dear Ms. Hedrick:

The Secretary has asked that I respond to your request for comments on the GAO draft report, "Congressional Oversight: Obtaining Information for Program Evaluation" (GAO Job Code 973766), which was transmitted to the Department of Education by your letter of September 2, 1994.

We believe it is very important for Congress to have timely accurate information regarding the operation of our programs and the report prepared by GAO will be helpful in furthering this aim.

We concur with your recommendation that a useful strategy for the Department and for Congressional staff is to increase dialogue with each other. Good communication, shared planning, and mutually agreed upon approaches are essential. Improved dialogue would not only strengthen the work that we do but ensure that we provide information Congress believes is important.

We also concur with the recommended strategy of having some core questions that will permit Congress and the Department to have information on programs during interim years. Our primary concern, however, is that the program goals, indicators, and outcome data requested for reauthorization be as consistent as possible with the requirements of the Government Performance and Results Act (GPRA) which is driving our strategic planning and budget processes.

We do have a concern with GAO's call for assessments "side-effects" and "comparative advantage". Questions such as these are often beyond the scope of the evaluation of particular programs, and difficult to address if the programs used for comparison are administered by another agency.

400 MARYLAND AVE., S.W. WASHINGTON, D.C. 20202-6100

Our mission is to ensure equal access to education and to promote educational excellence throughout the Nation.

Ms. Terry E. Hedrick
Page 2

Thank you for the opportunity to comment. My staff and I are prepared to respond if you or your representatives have any questions.

Sincerely,

Thomas W. Payzant

TWP:jt

cc: Alan Ginsburg

Comments From the Department of Health and Human Services

DEPARTMENT OF HEALTH & HUMAN SERVICES

Office of Inspector General

Washington, D.C. 20201

OCT 18 1994

Ms. Terry E. Hedrick
Assistant Comptroller General
United States General
Accounting Office
Washington, D.C. 20548

Dear Ms. Hedrick:

Enclosed are the Department's comments on your draft report, "Congressional Oversight: Obtaining Information for Program Evaluation." The comments represent the tentative position of the Department and are subject to reevaluation when the final version of this report is received.

The Department appreciates the opportunity to comment on this draft report before its publication.

Sincerely yours,

June Gibbs Brown
Inspector General

Enclosure

COMMENTS OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES ON THE GENERAL ACCOUNTING OFFICE (GAO) DRAFT REPORT "CONGRESSIONAL OVERSIGHT: OBTAINING INFORMATION FOR PROGRAM EVALUATION," SEPTEMBER 1994

The Department has reviewed the General Accounting Office (GAO) draft report and has the following comments.

MATTERS FOR CONGRESSIONAL CONSIDERATION

This report proposes a strategy to obtain information for program oversight and reauthorization that the Senate Committee on Labor and Human Resources may wish to adopt. The three components of this strategy are

-- selecting and adapting, from a core list, the descriptive and evaluative questions to be asked about a program in interim years and at reauthorization;

-- arranging explicitly to obtain timely oversight information in interim years as well as to receive results of evaluation studies at reauthorization; and

-- providing for increased communication with agency staff to help ensure that information needs are understood and that requests and reports are suitably framed and adapted as needs evolve.

This strategy can be adapted to take institutional realities into account. For example, in view of the many demands on its attention, the Committee might select future reauthorization questions for some programs and invite agencies to propose questions for others.

DEPARTMENT COMMENT

We generally support the GAO draft report's recommendation that congressional committees more clearly express their evaluative information needs to agencies and consult with agencies on those needs. However, we have concerns about GAO recommending a process that would burden agencies with additional evaluation reporting requirements on top of the extensive annual reporting of performance required under the recently enacted Government Performance and Results Act of 1993 (GPRA).

From our analysis, much of the information GAO cites as useful to congressional committees for both annual oversight and reauthorization activities could be part of the annual performance report required of each program activity under GPRA. The GAO draft report, however, does not give adequate recognition to this fact or to the point that Congress

2

recently addressed the need for regular evaluative information by enacting GPRA.

Major Contributors to This Report

Program Evaluation and Methodology Division

Stephanie Shipman, Assistant Director
Gail S. MacColl, Project Manager
Elaine Vaurio, Project Evaluator
Venkareddy Chennareddy, Referencer

Bibliography

References on Program Evaluation and Congressional Decision-Making

Abramson, Mark. The Funding of Social Knowledge Production and Application: A Survey of Federal Agencies. Washington, D.C.: National Academy of Sciences, 1978.

Barkdoll, Gerald L., and James B. Bell (eds.). Evaluation and the Federal Decision Maker. New Directions in Program Evaluation, No. 41. San Francisco: Jossey-Bass, 1989.

Bimber, Bruce. Congressional Support Agency Products and Services for Science and Technology Issues. Report prepared for the Carnegie Commission on Science, Technology, and Government, Sept. 17, 1990.

Boyer, John F., and Laura Langbein. "Factors Influencing the Use of Health Evaluation Research in Congress." Evaluation Review, 15:5 (Oct. 1991), 507-32.

Boruch, Robert, David Cordray, Georgine Pion, and Laura Leviton. "A Mandated Appraisal of Evaluation Practices: Digest of Recommendations to the Congress and the Department of Education." Educational Researcher, 10:4 (Apr. 1981), 10-31.

Bozeman, Barry, and Julia Melkers (eds.). Evaluating R&D Impacts: Methods and Practice. Boston: Kluwer Academic, 1993.

Caro, Francis G. "Leverage and Evaluation Effectiveness." Evaluation and Program Planning, 3 (1980), 83-89.

Chelimsky, Eleanor (ed.). "Program Evaluation: Patterns and Directions." American Society for Public Administration, 1985.

Citro, Constance F., and Eric A. Hanushek (eds.). Improving Information for Social Policy Decisions: The Uses of Microsimulation Modeling, Volume I, Review and Recommendations. Washington, D.C.: National Academy Press, 1991.

Cook, Thomas D., and William R. Shadish, Jr. "Metaevaluation: An Assessment of the Congressional Mandated Evaluation System for Community Mental Health Centers." In Gerald L. Stahler and William Tash (eds.), Innovative Approaches to Mental Health Evaluation, pp. 221-53. New York: Academic Press, 1982.

Cronbach, Lee J., and Associates. Toward Reform of Program Evaluation. San Francisco: Jossey-Bass, 1981.

Levitan, Sar A., and Gregory Wurzburg. Evaluating Federal Social Programs. Philadelphia: W.E. Upjohn Institute for Employment Research, Sept. 1979.

Lowi, Theodore J. "American Business, Public Policy, Case Studies, and Political Theory." World Politics, 16:4 (1964), 677-715.

Martin, Margaret E., and Miron L. Straf (eds.). Principles and Practices for a Federal Statistical Agency. Washington, D.C.: National Academy Press, 1992.

Mohr, Lawrence B. Impact Analysis for Program Evaluation. Newbury Park, Calif.: Sage Publications, 1992.

Rossi, Peter H., and Howard E. Freeman. Evaluation: A Systematic Approach, 4th Ed. Newbury Park, Calif.: Sage Publications, 1989.

Salamon, Lester M. (ed.). Beyond Privatization: The Tools of Government Action. Washington, D.C.: The Urban Institute Press, 1989.

Scriven, Michael. "Hard-Won Lessons in Program Evaluation." New Directions for Program Evaluation, 58 (Summer 1993).

Thompson, Mark. Decision Analysis for Program Evaluation. Cambridge, Mass.: Ballinger Publication, 1982.

Weiss, Carol H. Using Social Research in Public Policy Making. Lexington, Mass.: Lexington Books, 1977.

Wholey, Joseph S. Federal Evaluation Policy: Analyzing the Effects of Public Programs. Washington, D.C.: Urban Institute, 1970.

Windle, Charles, and J. Richard Woy. "From Programs to Systems: Implications for Program Evaluation Illustrated by the Community Mental Health Centers Program Experience." Evaluation and Program Planning, 6 (1983), 53-68.

Woods, Patricia. The Dynamics of Congress: A Guide to the People and Process in the Congress. Washington, D.C: Woods Institute, 1987.

References for Case Study Programs

Advisory Committee on Head Start Quality and Expansion. "Creating a 21st Century Head Start." Final Report. Washington, D.C.: Dec. 1993.

Arnold, Judith, Ann Zuvekas, and Nina Teicholz. "A Manual for Using Health Outcome Measures to Evaluate the Primary Care System." Discussion Draft, Feb. 1989. Washington, D.C.: Lewin/ICF, Task Order #340-86-0504.

Freeman, Howard E., K.J. Kiecolt, and Allan Harris, II. "Community Health Centers: An Initiative of Enduring Utility." Milbank Memorial Fund Quarterly: Health and Society, 60:2 (1982), 245-67.

Hawkins, Daniel, Jr., and Sara Rosenbaum. "Lives in the Balance: The Health Status of America's Medically Underserved Populations." Special report. Washington, D.C.: NACHC, Mar. 1993.

Puma, Michael J., et al. Prospects: The Congressionally Mandated Study of Educational Growth and Opportunity—The Interim Report. Washington, D.C.: U.S. Department of Education, 1993.

Stringfield, Sam, et al. Urban and Suburban/Rural Special Strategies for Educating Disadvantaged Children: First Year Report. Washington, D.C.: U.S. Department of Education, 1994.

U. S. Department of Education. Reinventing Chapter 1: The Current Chapter 1 Program and New Directions. Washington, D.C.: 1993.[1]

[1]This report drew on many other reports that were included in our review but are too numerous to be listed here. A full list of those reports is available from the Department of Education's Planning and Evaluation Service.

Related GAO Products

Elementary School Children: Many Change Schools Frequently, Harming Their Education (GAO/HEHS-94-45, Feb. 4, 1994).

Public Health Service: Evaluation Set-Aside Has Not Realized Its Potential to Inform the Congress (GAO/PEMD-93-13, Apr. 8, 1993).

Exiting Program Improvement (GAO/HRD-93-2R, Mar. 30, 1993).

Chapter 1 Accountability: Greater Focus on Program Goals Needed (GAO/HRD-93-69, Mar. 29, 1993).

Improving Government: Measuring Performance and Acting on Proposals for Change (GAO/T-GGD-93-14, Mar. 23 1993).

Compensatory Education: Difficulties in Measuring Comparability of Resources Within School Districts (GAO/HRD-93-37, Mar. 11, 1993).

Improving Government: Need to Reexamine Organization and Performance (GAO/T-GGD-93-9, Mar. 11, 1993).

Compensatory Education: Additional Funds Help More Private School Students Receive Chapter 1 Services (GAO/HRD-93-65, Feb. 26, 1993).

Government Management Issues (GAO/OCG-93-3TR, Dec. 1992).

Children's Programs: A Comparative Evaluation Framework and Five Illustrations (GAO/PEMD-88-28BR, Aug. 31, 1988).

United States
General Accounting Office
Washington, D.C. 20548-0001

Bulk Mail
Postage & Fees Pai
GAO
Permit No. G100

Official Business
Penalty for Private Use $300

Address Correction Requested

Ordering Information

The first copy of each GAO report and testimony is free. Additional copies are $2 each. Orders should be sent to the following address, accompanied by a check or money order made out to the Superintendent of Documents, when necessary. Orders for 100 or more copies to be mailed to a single address are discounted 25 percent.

Orders by mail:

U.S. General Accounting Office
P.O. Box 6015
Gaithersburg, MD 20884-6015

or visit:

Room 1100
700 4th St. NW (corner of 4th and G Sts. NW)
U.S. General Accounting Office
Washington, DC

Orders may also be placed by calling (202) 512-6000 or by using fax number (301) 258-4066, or TDD (301) 413-0006.

Each day, GAO issues a list of newly available reports and testimony. To receive facsimile copies of the daily list or any list from the past 30 days, please call (301) 258-4097 using a touchtone phone. A recorded menu will provide information on how to obtain these lists.

PRINTED ON RECYCLED PAPER