1976

Evaluation of a pilot family planning project in the Iowa Cooperative Extension program

Eva Christina Elisabeth Aspegren
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Home Economics Commons

Recommended Citation
Aspegren, Eva Christina Elisabeth, "Evaluation of a pilot family planning project in the Iowa Cooperative Extension program " (1976). Retrospective Theses and Dissertations. 5675.
https://lib.dr.iastate.edu/rtd/5675

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in “sectioning” the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from “photographs” if essential to the understanding of the dissertation. Silver prints of “photographs” may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106
ASPEGREN, Eva Christina Elisabeth, 1933-
EVALUATION OF A PILOT FAMILY PLANNING PROJECT
IN THE IOWA COOPERATIVE EXTENSION PROGRAM.

Iowa State University, Ph.D., 1976
Education, home economics

Xerox University Microfilms, Ann Arbor, Michigan 48106
Evaluation of a pilot family planning project
in the Iowa Cooperative Extension program

by

Eva Christina Elisabeth Aspegren

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major: Home Economics Education

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa
1976
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>INTRODUCTION</strong></td>
<td>1</td>
</tr>
<tr>
<td><strong>LITERATURE REVIEW</strong></td>
<td>9</td>
</tr>
<tr>
<td>Program Evaluation</td>
<td>9</td>
</tr>
<tr>
<td>The professional judgment approach</td>
<td>10</td>
</tr>
<tr>
<td>The measurement approach</td>
<td>11</td>
</tr>
<tr>
<td>The Tyler evaluation model</td>
<td>12</td>
</tr>
<tr>
<td>Evaluation as input for decision-making</td>
<td>14</td>
</tr>
<tr>
<td>Broad-aimed programs and evaluation</td>
<td>63</td>
</tr>
<tr>
<td>Innovative programs and evaluation</td>
<td>64</td>
</tr>
<tr>
<td>Issues and problems in program evaluation</td>
<td>70</td>
</tr>
<tr>
<td>Evaluation of Extension and Adult Education Programs</td>
<td>93</td>
</tr>
<tr>
<td>Evaluation of programs aimed at low-income adults</td>
<td>95</td>
</tr>
<tr>
<td>Evaluation of Family Planning Programs</td>
<td>116</td>
</tr>
<tr>
<td>KAP studies</td>
<td>117</td>
</tr>
<tr>
<td>Adoption of contraceptives</td>
<td>119</td>
</tr>
<tr>
<td>Decline in fertility</td>
<td>121</td>
</tr>
<tr>
<td>Program efficiency</td>
<td>121</td>
</tr>
<tr>
<td>Theoretical issues</td>
<td>121</td>
</tr>
<tr>
<td>Systems for evaluating national programs</td>
<td>123</td>
</tr>
<tr>
<td>Review of evaluation studies</td>
<td>124</td>
</tr>
<tr>
<td>Family planning and low-income audiences</td>
<td>128</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>135</td>
</tr>
<tr>
<td><strong>METHODOLOGY</strong></td>
<td>137</td>
</tr>
<tr>
<td>Project Under Study</td>
<td>137</td>
</tr>
<tr>
<td>Theoretical Framework</td>
<td>137</td>
</tr>
<tr>
<td>Participants in the Study</td>
<td>139</td>
</tr>
<tr>
<td>Formulation of Questions to be Answered by the Study</td>
<td>141</td>
</tr>
<tr>
<td>Context</td>
<td>144</td>
</tr>
<tr>
<td>Input</td>
<td>144</td>
</tr>
<tr>
<td>Process</td>
<td>145</td>
</tr>
<tr>
<td>Product</td>
<td>147</td>
</tr>
<tr>
<td>Topic</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Data Collection</td>
<td>147</td>
</tr>
<tr>
<td>Treatment of Data</td>
<td>150</td>
</tr>
<tr>
<td>FINDINGS AND DISCUSSION</td>
<td>151</td>
</tr>
<tr>
<td>Context Evaluation</td>
<td>154</td>
</tr>
<tr>
<td>Theoretical information</td>
<td>155</td>
</tr>
<tr>
<td>Empirical information</td>
<td>159</td>
</tr>
<tr>
<td>Stated objectives</td>
<td>160</td>
</tr>
<tr>
<td>Summary</td>
<td>161</td>
</tr>
<tr>
<td>Input Evaluation</td>
<td>162</td>
</tr>
<tr>
<td>Identified resources and obstacles</td>
<td>163</td>
</tr>
<tr>
<td>Selected strategies</td>
<td>166</td>
</tr>
<tr>
<td>Process Evaluation</td>
<td>170</td>
</tr>
<tr>
<td>Changes during the implementation of the pilot project</td>
<td>172</td>
</tr>
<tr>
<td>Leadership</td>
<td>180</td>
</tr>
<tr>
<td>Understanding and agreement with program intent</td>
<td>182</td>
</tr>
<tr>
<td>Training methods</td>
<td>186</td>
</tr>
<tr>
<td>Training of state specialists and local home economists</td>
<td>188</td>
</tr>
<tr>
<td>Approaches used by paraprofessionals</td>
<td>189</td>
</tr>
<tr>
<td>Product Evaluation</td>
<td>190</td>
</tr>
<tr>
<td>Accomplishments of the pilot project as related to objectives</td>
<td>193</td>
</tr>
<tr>
<td>Information needed for recycling decisions</td>
<td>198</td>
</tr>
<tr>
<td>SUMMARY, CONCLUSIONS AND RECOMMENDATIONS</td>
<td>202</td>
</tr>
<tr>
<td>Summary</td>
<td>202</td>
</tr>
<tr>
<td>Conclusions and Recommendations</td>
<td>206</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>211</td>
</tr>
<tr>
<td>REFERENCE NOTES</td>
<td>225</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>226</td>
</tr>
<tr>
<td>APPENDIX</td>
<td>227</td>
</tr>
<tr>
<td>Questions that Might be Answered Through the Proposed Study</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Context</td>
<td>227</td>
</tr>
<tr>
<td>Input</td>
<td>227</td>
</tr>
<tr>
<td>Process</td>
<td>228</td>
</tr>
<tr>
<td>Product</td>
<td>229</td>
</tr>
</tbody>
</table>
INTRODUCTION

Education today is challenged to meet the rapidly shifting social, economic, and intellectual needs of society. At the same time, there is a growing call for accountability of education to the individual as well as to society. Much of this trend is caused by legal requirements to evaluate programs supported with federal and state funds, but the quest goes beyond budgetary aspects. The diversity of educational values, needs, and expectations by students or clients, parents or citizens, and educators have forced the issues, especially since the federal government and some states have established broad educational improvement programs, aimed at social action, ultimately to improve the quality of life for the recipients.

Educational efforts have always been evaluated. Judgments, based on personal observations have ubiquitously been, and continue to be passed by people involved at all and any levels. Verdicts by people with expertise and experience, or professional judgment, are still being used as a method of evaluation, especially within higher education. The realization that the scientific method of inquiry, i.e., the identification of a problem, gathering of relevant data and testing of the hypothesis empirically, made possible a more rapid advancement in academic disciplines caused the measurement approach to evaluation to emerge around the turn of this century. In the 1930's, the Tyler evaluation model introduced educational objectives and the assessment of congruence between stated objectives and educational achievement of either an individual or a program.
The measurement approach focuses on the outcome of the educational process, or the end product, as does the Tyler model, although in the latter the purpose of evaluation was broadened to include such things as periodic checks on the effectiveness of the institution and helping teachers and students to clarify their purposes and to see more concretely the directions in which they are moving. In education it often takes a long time to ascertain if the end product is acceptable. In addition, this kind of evaluation rarely indicates where changes are needed within the program. Adjustments to changing external and internal situations need to be made at an early stage in order to maintain desired outcomes.

The 1960's have seen a trend in evaluation to provide a data base for judgments about program and project decisions. The approach has a clear management and systems orientation. Any educational program, aspiring to meet the needs of society, must consciously direct its resources and efforts toward the attainment of these goals. In this process, the program planner is faced with the constant task of decision-making. Because any kind of decision involves a choice situation, the decision-maker makes a conscious or unconscious evaluation of available options. Ideally, decisions should be based on sound information about alternative plans and procedures, and not on intuition or limited knowledge. Thus evaluation is faced with the task to delineate, obtain and provide useful information for judging decision alternatives, a definition used in the Context, Input, Process and Product (CIPP) evaluation model.
The Cooperative Extension Service is an educational program through which knowledge is transmitted from the researcher to the people. Traditionally, its programs have centered around agriculture, home economics, and youth work, mainly reaching the middle socio-economic and rural strata of society. Today, the trend is to broaden the scope of program activities and to reach a wider audience with the ultimate aim of developing community resources and improving quality of life.

One effort to reach a wider audience is through the Expanded Nutrition Program. The aim of the program is to improve nutritional intake and practices of the families through education, thus improving their level of living. The extension educator is generally viewed as a change agent who creates learning situations and serves as a resource person to his/her clients in problem solving. The ENP program uses paraprofessional aides, recruited from the neighborhood to reach the low-income clients as frequently the characteristics, values, and attitudes of these families prevent them from seeking and accepting knowledge and advice from the middle-class oriented extension worker.

Although the aides from the start of the ENP program were to teach nutrition and food preparation skills only, it was soon found that they had to broaden their teaching in order to improve the level of living of their clients. Thus they had to be prepared to answer questions and to make suggestions about sanitation, money management, food buying, and food storage.

A nationwide survey (Brand, 1971) revealed that nearly two-thirds of the paraprofessionals were being asked for family planning or birth control advice by the women and girls in the program. Moreover, most program
aides saw a need for family planning education among the clients, strongly favored family planning and birth control, and would be willing to teach homemakers about these subjects if they themselves were trained to do so.

Concern for the increase in world population has been voiced for centuries, but not until the last decade has the "population explosion" become an urgent public issue. There is a dawning realization that the resources on earth are finite, and that it takes a shorter and shorter time for the world population to double. The consequences of too rapid a population growth on a worldwide basis are problems in securing adequate nutrition for the expanding population as well as in raising its living standard, especially for the one-half to two-thirds of the world that already lives in deep poverty.

Based on current thinking that man can control his future population through his own effort and that family planning is one means to do so, organized efforts to regulate population growth have been instigated. In the Western world the movement frequently started through the endeavor of a few individuals, spread through private organizations, and from there gained national attention. In other parts of the world, the pattern has often been the reverse. National family planning programs, or large scale efforts to provide services to individuals which enable them to practice family planning, did not become prominent until the 1960's. One contributing factor might have been that, since 1965, the United States government began to advocate family planning as one part of its foreign technical assistance program.

The United Nations has declared that all persons have the right to decide freely and responsibly on the number and spacing of their children.
The world organization proclaimed 1974 as the World Population Year, which culminated in a Population Conference in Bucharest, Rumania. The outcome of this world meeting illustrated the controversial nature of the concept of family planning. It is an euphemism that is rather ambiguous in its real meaning. For some people, family planning refers to specific techniques of contraception, for others it relates to very broad programs for improving the physical and social well-being of families. It is associated with assisting individuals in having the number of children they want, when they want; but also with imposed fertility limitations. Family planning is often used synonymously with birth prevention, birth control, and planned parenthood. Several international conferences and seminars, dealing with the assessment of national family planning programs, have spent much time in trying to define terms in order to make a meaningful evaluation possible. Examples of such conferences are the International Conference on Family Planning Programs, sponsored by the Ford Foundation, the Population Council and the Rockefeller Foundation, and held in Geneva, Switzerland in 1965; a Regional Seminar on Evaluation of Family Planning Programmes, sponsored by the Economic Commission for Asia and the Far East, and held in Bangkok, Thailand in 1969; and the Fourth Annual Population Conference, sponsored by the Organization for Economic Co-operation and Development, and held in Paris, France, in 1971.

Up to the 1960's publicly the United States was opposed to birth control, some states preventing the sale of contraceptives and the dissemination of information about them through legislation. Because of the evolving public concern about population, energy, and environment, this stand has changed and family planning has become incorporated into federal
social action programs. The War on Poverty program, aimed at the low-income and minority people, was one of the first to include some family planning efforts, and thereafter the Office of Economic Opportunity has added family planning to its programs. Since early 1970's, federal funds have been made available for low-income family planning programs and research in this area. The goal of the federal government for 1971-1972 was to reach 30 percent or 1.7 million of all the low-income fertile families in the U.S., a figure that by 1975 would have increased to 6.6 million families (Rogers, 1973).

The project in this study was an outgrowth of an International Family Planning Workshop held at Iowa State University in 1973 and sponsored by the American Home Economics Association (AHEA) and the College of Home Economics. It was believed that in order to train personnel and develop materials to be used internationally, methods, techniques, and strategies needed to be tried out in the United States under conditions as similar as possible, i.e., in extension and/or adult education, and cross-culturally. In particular, the effectiveness of integrating family decision-making with home economics subject matter and family planning education was of special interest. The term family planning refers to the decision by individuals to have the number of children they want and when they want.

Therefore, a pilot project within the ENP in southeast Iowa was established. The present study deals with the evaluation of the first phase of the program. As evaluation ultimately aims at improving programs, it plays an important role in any program, but becomes crucial in
a pilot program where many decisions have to be made without support of prior evidence as well as when recycling decisions are made.

The first proposal for the project, dated November, 1973, was drafted jointly by Cooperative Extension Home Economics and the College of Home Economics, Iowa State University. Once funded, an assistant state leader assumed the responsibility for the development and execution of the project. During the data gathering period, a state level planning committee was established, local home economics staff and paraprofessional aides received training, aides started to work with clients, and visual materials to be used by the aides were developed. Some internal effort to assess training needs on part of the local staff and aides was made by the assistant state leader, as was an evaluation of the entire project.

An extensive search of literature was made in order to find a theoretical framework for evaluation of a pilot program, where the communicated message carries a taboo stigma and in some instances has a negative significance, and where the audience necessitates unobtrusive, nonthreatening techniques of assessment. Furthermore, as the pilot project was a single project, involving few people, there was a need to document findings from the literature.

The objectives of the study were to record and describe the planning and implementation of a pilot program; to analyze the processes and procedures used in the program; and to make recommendations about processes and procedures, based on the analysis, for other similar programs.

Assumptions underlying the study were that subjects involved in the program planning and execution were willing and able to provide needed information; that the Context, Input, Process and Product (CIPP)
evaluation model could be entered at any stage; and that information gained from a pilot program can be useful in other similar programs.

Several limitations surround the study. It was limited to one program only. Because of the limited time the investigator could participate in the pilot program, major emphasis was on the process evaluation, which in turn was based on periodic written and oral communications instead of day-by-day observations. Neither program participants nor the paraprofessional aides were studied in any depth due to the relatively short time-span of the study and to the initial stage of the project. The number of subjects in the study was small, hence statistical analysis was limited and inferences outside the specific program are very tentative.
LITERATURE REVIEW

Program Evaluation

Currently, there is an emphasis on evaluation and accountability within education and social action programs. Much of it is caused by legal requirements to evaluate programs supported with federal and state funds, but the trend for accountability goes beyond budgetary aspects. Social and humanitarian factors enter as accountability to the individual and to society. Jacobs (1974) points out that the diversity of educational values, needs, and expectations by students, parents, and educators causes both the profession and its clients to seek and often demand alternatives in both the means and ends of education. The same tendency is present in social action programs.

Policy-makers increasingly feel a need to rely less on subjective and political reasons for making decisions, and are asking for hard data on which to base more rational decisions (Evans, 1974). On the federal level, the introduction of the Program Planning and Budgeting System brought out the need for data on program effectiveness in terms of objectives, methods, and outcomes alongside the standard administrative budget considerations.

Prior to 1964, objects of evaluation in education consisted almost exclusively of small programs concerned with such things as curriculum development or teacher training. They generally occurred in a single school or school district and sought to produce educational change on a limited scale. Typically they involved modest budgets and small research staffs (Cohen, 1970). This changed when the federal government and some
states established broad educational improvement programs, aimed at social action, directed at millions of children, created by Congress and administered by federal agencies far from the school districts which actually designed and conducted the individual projects.

Program evaluation is becoming more feasible than previously because funds are increasingly being allocated for this purpose and because sophisticated methodologies and technologies are being developed. There is also an increased awareness on the part of administrators and policy-makers of the utility of evaluation research and of the benefit of linking the evaluative function to the decision-making structure within an organization or agency.

It is assumed that program evaluation ultimately will lead to improvements of present or future programs. As Suchman (1969) indicates, through an evaluation study it is possible to analyze a program in terms of its objectives, the assumptions underlying those objectives, the specific program activities designed to achieve those objectives, and the rationale for believing that activities are capable of attaining the objective. Evaluation may also separate the "idea" of the program from how it is being carried out and determine criteria for observing the extent to which the objectives are being attained.

The professional judgment approach

Educational efforts have always been evaluated; judgments, based on personal observations have ubiquitously been passed by people involved at all levels, e.g., students, instructors, administrators, communities. The thought that evaluation equals judgment still exists today, with the implication that the verdict is passed by people with expertise and
According to the Phi Delta Kappa (PDK) Committee (1971), the professional judgment approach commonly occurs within higher education at the present time. The accrediting associations' visiting teams use this concept of evaluation. Although data-gathering is done by the school itself, the formal discernment is passed by the visitation team.

The professional judgment approach has the advantages of being a reasonably quick procedure as well as one that takes the unique features of the program into consideration. The main drawback is that because of questionable reliability and objectivity, generalization of findings is difficult. Data and criteria used are ambiguous and hidden in the process itself.

The measurement approach

The scientific method of inquiry - identification of a problem, gathering of relevant data and testing the hypothesis empirically - has been employed increasingly in various academic disciplines in order to facilitate more rapid advancement in their fields. Education, although an applied science, realized that intuition, however valuable in itself, does not provide a sound basis for progression. Formal assessments of educational efforts started around the turn of this century and evaluation became identical to measurement as reported by the PDK Committee (1971). Emphasis was on empirical research (Suppes, 1974), which made mathematical and statistical manipulation and analysis possible. Student performance was measured by means of standardized tests, developed by measurement specialists. Standardized tests provide norm-referenced information about student achievement, hence the student (class or school) can be compared
with other students (classes or schools) and data interpreted accordingly. These group norms are demographic in nature and do not take either social or cultural factors into consideration; nor do they leave any room for value judgments, as objectivity is stressed. Evaluation became restricted to the limited number of variables for which measuring instruments could be developed. Variables which could not be measured were largely ignored and became ultimately unimportant. All in all, this approach to evaluation is very narrow, mechanical and instrumentally focused.

**The Tyler evaluation model**

Another approach, originating during the 1930's, introduced educational objectives and the assessment of congruence between stated objectives and educational achievement of either an individual or a program. The name of Ralph W. Tyler is associated with this evaluation model. Focus is still on terminal concerns or end products, but the purpose of evaluation was broadened to include periodic checking on the effectiveness of the educational institution; validating the hypothesis upon which the institution operates; providing information basic to effective guidance of individual students; providing psychological security to school, staff, parents, and students; providing a sound basis for public relations; and helping teachers and pupils to clarify their purposes and to see more concretely the directions in which they are moving.

The evaluation design commonly used in the Tyler model is borrowed from the psychological laboratory research model and consists of pretest, treatment and one or more posttests. Control groups are used, efforts are made to keep all but the independent variable constant, and objectivity is strived for (Sjogren, 1974).
The Tyler evaluation model includes deciding in advance what the final outcome ought to look like, and then using this conception to prescribe methods for attaining and measuring it (Clark, 1975). In 1949, Tyler identified the steps as the formulation of objectives and classification of these according to level of specificity; the definition of each objective in terms of student behavior; the identification of situations in which students can be expected to display these types of behavior; development or selection of techniques for appraising student behavior; and the gathering and interpreting of data.

Several assumptions underlay this evaluation model, i.e., that education seeks to change behavior; and that educational programs are appraised by finding out how far these objectives are actually being realized. Other assumptions are: that any device which provides valid evidence regarding progress of students towards educational objectives is appropriate; and that the participation of teachers, pupils, and parents in the process of evaluation is essential to derive maximum value.

The Tyler model focuses on objectives, which has connotation for designing instructional strategies as well as for providing opportunities for feedback and process data. Objectives are at different levels. The broad and program objectives are usually selected by school administrators and teachers, based on need assessment and screened through the philosophies of the school and educators and learning psychology (Guba, 1969; PDK Committee, 1971; Tyler, 1949). In order to prescribe learning experiences, select materials, teaching strategies and sequence, as well as making the assessment of progress towards the objectives feasible, the
broad objectives need to be narrowed down and stated in behavioral form, preferably in student behavior.

The more specific the behavior is stated, the easier it is to access progress, but at the same time, evaluation is confined to these predeter-
mined, specific behaviors. As Clark (1975, p. 29) points out, "Neither educator nor evaluator is presently skillful enough to predetermine optimal and complete sets of critical properties for many of our crucial educational objectives." Education is too complex a process to be measured by only overt, specific, predetermined behavior (Weiss, 1972a; Clark, 1975; Stufflebeam, 1971).

The emphasis on student behavior has also stressed the evaluation of the product, the outcome of the educational process, and at the same time diminished the importance of evaluating the process itself. Another drawback is that whatever process evaluation is done, the result is not utilized in the ongoing program from which the data were extracted, but in the following ones (Weiss, 1972c; Clark, 1975; PDK Committee, 1971). This development has come about as the evaluation model has been applied. As mentioned before, the Tyler model in its theoretical form provides for both feedback and process data. Nor did Tyler visualize that objectives were to become so specific, either (Shane & Shane, 1973).

Evaluation as input for decision-making

During the 1960's a third trend in evaluation emerged, namely evaluation as input for decision-making. Basically, evaluation in this sense provides a data base for judgments which lead to program and project dec-
cisions. The approach has a clear management and systems orientation.
Systems theory Although systems theory originated within the industrial world, it was soon found to be a useful analytic framework when applied to organizations. During the last 20 years, systems thinking has had its greatest period of development and is increasingly employed within the social and behavioral sciences.

The systems movement is not a theory in the pure sense of the word (Immegart & Pilecki, 1973) but rather a systematic, relational mode of thought. It is concerned with the totality of behavior, or function, in context and within the chronological development.

A system can be viewed as a combination of interdependent elements and their activities which transfer inputs to outputs. Key aspects of a system are that it is composed of various parts or elements or functions; that the parts are interrelated, e.g., the behavior of each element is dependent on the behavior of the other elements; that there is "feedback" between the parts, positive or negative, facilitating control of the system; that there are inputs and outputs; and that it has boundaries, separating it from other systems.

An educational or a social action program or organization, from a systems point of view, can be considered as a multipurpose, adaptive, and open system, e.g., one that has multiple interactions with a dynamic environment and that is able to receive and respond to input from the same in a way that is desirable according to the system's objectives (Shein, 1970). The dynamic environment places demands and constraints upon a system in various ways, hence the total functioning of a system cannot be
understood without explicit consideration of these environmental demands and constraints.

In most instances, it is difficult to specify clearly the boundaries of a particular social system, as it has multiple links, or interfaces, with its environment. Furthermore, all systems have subsystems, which in turn may have subsystems. Any subsystem is a system in and of itself. At the same time, all systems are also subsystems to larger and more complex systems. Hence system boundaries are ultimately dictated by utility or feasibility.

Feedback, positive or negative, is information regarding outputs re-entering the system as input to affect succeeding outputs (Deacon & Firebaugh, 1975). Within an organization, this becomes a communication process. To the cybernetic school of systems thinking, all systems can best be understood through their communication and control activities (Immegart & Pilecki, 1973). To provide for necessary flow of information, adequate and functioning channels of communication become vital.

Systems control theory assumes that output is determined by inputs, some of which can be controlled, and by the internal state of the system (Zemach, 1973). A controller is a component of the system which can in some way manipulate the controllable inputs, including policies, so as to improve the behavior of the system, according to the desired output. This controller can be identified with the administrative or decision-making function. Action taken is partly or wholly based on feedback about system performance.
Feedback control introduces corrective action to counteract disturbances to the system when effects of disturbance become evident in the system or its output. Information, gained through the use of sensor elements outside the system boundaries which perceive relevant environmental changes, can also be fed forward and corrective manipulation introduced before errors become evident.

Evaluation of a program within the systems control theory becomes an integral part of the process of information feedback and feed forward, and comparison to desired conditions. This implies that evaluation must be an ongoing process within the system, linked to the decision-making levels, and employing both feedback and feed forward information. The process needs to be planned along with the planning of the program itself, according to existing requirements and constraints.

**Systems analysis** Systems analysis is built on general systems theory and decision-making and is a technique used in program planning. Evaluation is involved in the process. Several procedures have evolved and have been applied within education, i.e., the Critical Path Method, the Program Evaluation and Review Technique, the Cost-Benefit Analysis, and the Planning, Programming, and Budgeting System.

**The Critical Path Method** The Critical Path Method (CPM) identifies critical events, that in turn might require subevents. Their interrelationships are diagnosed. As a program is conducted, information about events is obtained and fed back into the system in order to modify future events (Knowles, 1970).
The Program Evaluation and Review Technique (PERT) is a refinement of the Critical Path Method. Every event that must take place in order for the ultimate goal to be achieved is identified. A flowchart, or network, depicting the interdependence and sequence of events is drawn. Series of activities that can be performed concurrently are drawn in parallel paths. Time required for each activity is estimated by using a formula, which takes into consideration three estimates: optimistic time, most likely time, and pessimistic time. The path requiring the largest time becomes the critical path, other paths are slack paths, leaving lax time (Knowles, 1970).

In itself, PERT solves no problems, but through its application one is able to gather the information necessary for efficient appropriation of human and material resources, to detect obstructions in achieving desired goals, and to estimate project and sub-project time requirements. It is also a means of reviewing the logical steps and their interrelationships in achieving a desired outcome.

For a project with a reasonable number of events, a paper-and-pencil PERT is possible, but for more complex and larger projects, access to a computer becomes imperative (Immegart & Pilecki, 1973).

The Cost-Benefit Analysis adds another dimension to the system analysis. It provides the decision-maker with a means of assessing the costs, benefits, and consequences of all alternatives in a program at a given time. The ratio of benefits to costs is an indication of the return from the investment in the program. The analysis introduces elements of clarity, comparability, and simplification.
into complex situations. Emphasis is on planning, but the method can also be applied retrospectively to calculate the returns on investment in past programs. In both instances, evaluation aspects are involved, i.e., what kind of and how many benefits.

The Cost-Benefit Analysis attempts to identify the benefits of a program, both tangible and intangible, to look at the direct and indirect costs of conducting the program, and to put them into a common unit of measure, dollars (Weiss, 1972a).

There are several problems inherent in this method: the intangible benefits of a social action or educational program can be hard to quantify, discerning indicators of benefits have to be selected, and factors not susceptible to monetary expression have to be reported quantitatively. The distributional effects of costs and benefits within a program and their relative values are not reflected in the total dollar calculation, but may well be vital for policy-making (Rossi, 1972; Weiss, 1972a). Another question is what factors to include in the calculation of benefits. A difference in assumptions can result in greatly diverse cost-benefit conclusions in the same program. To assign a dollar value to a benefit is another hurdle in this method, as is the projection of dollar value of future benefits.

The Cost-Benefit Analysis is often a part of other evaluation methods. Weiss (1972a) considers it most useful when: (1) existing data (or easily collectible or credibly reconstructable data) indicate the scope and degree of program impact; (2) when the main benefits can be reduced to dollar terms without overly fancy guess work or neglect of
crucial effects; and (3) when general benefit level rather than the distributional change of benefits is the main criterion.

The Planning, Programming and Budgeting System

As a further refinement of the systems approach, the procedures of the Planning, Programming and Budgeting System (PPBS) was developed by the Department of Defense in the early 1960's and became widely used when President Lyndon Johnson directed all the major civilian agencies of the federal government to install the system and to establish a central program policy and planning staff to assist in its implementation (Kirst, 1975). By using the PPBS, an agency would be able to identify national goals with precision and on a continuing basis and to choose among those goals the ones that are most urgent, as well as to search for alternative means of reaching those goals most effectively at the least cost. It could also inform on cost on a long-range basis, and measure program performance to insure a dollar's worth of service for each dollar spent (Orlans, 1973).

Within the concept of PPBS, analysis becomes a critical factor in the decision-making process. The analysis, however, is done before the fact, as opposed to after-the-fact, and implies continual evaluation and review of the organizational goals and financial strategy for realizing those goals (Immegart & Pilecki, 1973).

The PPBS serves primarily to facilitate organizational decision-making at the administrative levels. The ultimate thrust of the PPBS is economic rationality, i.e., measuring the cost-effectiveness ratio of alternative lines of action and tying the budgeting process directly to the planning process (Knowles, 1970).
The PPBS has been used both in social action and education as an answer to the quest for systematic analysis and accountability. In the last few years, its popularity has decreased. Like any method, PPBS has disadvantages as well as advantages. Even if viewed as a purely technical issue that can be handled by an accounting firm, PPBS is a complex, time-consuming process. The main difficulty is that it requires elaborate and sophisticated analysis, which needs to be based on sufficient and adequate data in order to be meaningful. Without these kinds of data, the procedure becomes a series of guesses. The objectives of the organization need to be explicit, operational, and stated in terms of outcomes that can be measured against criterion standards. Here the issue becomes not merely technical, but also value laden. The number of objectives tends to rapidly become unwieldy.

Unless specially provided for, PPBS does not leave room for emergencies or new activities. Furthermore, it assumes rationality, which may conflict with political activities surrounding the organization. A comprehensive, effective communication system is required throughout the organization with maximal speed, feedback, and capacity for information transmission.

On the whole, PPBS in its present form or with modifications has potential for linking evaluative data to the decision making process.

**Decision-making**

Decision-making can be defined as the making of choices or the resolving of alternatives (Deacon & Firebaugh, 1975). In itself, decision-making does not imply action as does problem-solving, but is part of the problem-solving and management processes.
Whenever decision-making is discussed, it is usually assumed it refers to the rational model, which is built on the rational man theory. The latter focuses on consistent, programmable, organized thinking behavior of man, and gives primacy to behavior that is goal related and is consistent with present beliefs. It excludes feelings and emotions of man (Argyris, 1973). But people may do things without fully understanding why. Decisions can be made intuitively, yielding an immediate decision, bypassing well-defined, conscious steps. Intuition is aided by experience, but does not necessarily result from it. Attention is on a total conception of the problem, elements in their own rights are neglected. All or parts of decisions may be intuitive (Deacon & Firebaugh, 1975).

A second procedure for choice, that does not include rationality, is the process of tradition and faith, through which people do things because that is the way they are or always have been done (Argyris, 1973). Man's capacity to process information and solve problems is finite. Because the environment in which man finds himself is so complex, man cannot make decisions without having constraints imposed upon his environment so as to make it manageable. Hence it is not possible for man to make a totally rational decision as all possible alternatives cannot be identified.

Rational decision-making is a hierarchical process which includes becoming aware that a decision is needed, identifying and weighting appropriate alternatives, and choosing among or resolving alternatives.

Identification of unmet needs and unsolved problems as well as actively searching for opportunities outside the immediate environment can create awareness of need for decision. The creation of awareness that
decisions need to be made is an essential and continuing function of any evaluation program (PDK Committee, 1971). The decision situation needs to be stated, preferably in question form, and the authority and responsibility for making the decision identified, e.g., the loci of decision-making and veto power. Decision alternatives do not just emerge; they must be discovered or invented and then specified in operational terms so that they can be assessed. The number of alternatives considered in decision making varies with the importance of the decision to be made or with the individual's capacity to generate alternatives (Deacon & Firebaugh, 1975).

Once the explicit and/or implicit alternatives for choice-making are defined and formulated, criteria for weighing the decision alternatives are determined, and decision rules established. Criteria may be looked upon as yardsticks for values, as providing a means of measuring variables to see how far it matches the ideal, or the value. The criteria used by a decision-maker are formed through an interaction of his knowledge and values (PDK Committee, 1971).

The values underlying the criteria are those of the decision-maker and his or her values, to some extent, reflect those held about the external and internal environment. At least four kinds of values encroach on an educational criterion: institutional values, values external to the institution, subsystem maintenance values, and private values. In a value conflict, an experienced decision-maker will attempt to reflect consensus of the competing values to a degree that will gain adequate support for the decision. Rational decision-making assumes objectivity in weighing
alternatives, taking such factors into account as the total set of environmental circumstances, the urgency required, the ways alternatives and/or criteria were generated, the predictability of consequences for each alternative, and the cost and risks associated with these. Decision rules specify the conditions under which any decision would be chosen over all other alternatives under consideration (Deacon & Firebaugh, 1975). These rules may be economical in nature, resulting in the cheapest alternative that will do the job at an acceptable level; or optimal, searching for the best return for expanded resources. Satisficing implies that the search for alternatives ends when one is found good enough to meet the anticipated outcome.

Decisions can be classified into programmed and nonprogrammed ones. Programmed decisions are prescribed. Before a choice situation occurs, a definite procedure has been worked out for handling them through established laws, rules, regulations, and organizational policy. Limits and boundaries are well defined. Programmed decisions are repetitive and routine. A high percentage of educational decision making falls into this category. Nonprogrammed decisions involve a number of unknown variables, novel, unstructured, and consequential (Immegart & Pilecki, 1973). Rational man deals with these decisions by orderly thinking, excluding intuition, spontaneity, and faith (Argyris, 1973). Argyris makes the observation that programmed activity tends to drive out nonprogrammed activity within an organization.

Information is needed for all stages in rational decision-making. Decisions can be made with certainty, that is with complete knowledge of
consequences of the decision, or with varying degrees of uncertainty, involving risk-taking. Risks in decision-making, or the chance of losses, can be known or unknown. Known risks are based on knowledge of objective probability, based on experience, repeated experiments, or outcome of large samples; or subjective probability, based on the decision-maker's belief in the outcome (Deacon & Firebaugh, 1975).

Schmidtlein (1974) discusses two competing, ideal, paradigms for the policy-making decision-making processes that occur in educational operations: the comprehensive/prescriptive (C/P), and the incremental/remedial (I/R). The C/P is typified by the concept of planning, as in the PPBS approach, systems analysis. It is the dominant orientation of today's educational policy-makers, especially at the federal level. The I/R is exemplified in the political decision-making process underlying the Context-Input-Process-Product (CIPP) evaluation model, and is more compatible with the dominant philosophical orientation of the American society.

The C/P and I/R paradigms conflict in areas of legitimacy and effectiveness; legitimacy because of fears that the C/P might lead to control and centralized authority, and the I/R to power accumulation in a few hands not accountable to the public. The conflict assumes that each paradigm is feasible in all areas of policy-making and their constraints lie in the characteristics of people. Schmidtlein (1974) takes the position that the paradigms are not equally effective in all environments, as constraints in the environment of the area subject to decision limit the usefulness of each paradigm. These constraints can be classified into those imposed by time, resources, and function, i.e., lack of time,
availability and distribution of resources, degree of consensus in the policy field, and conflicting functional demands.

In education there exists belief in the efficiency of the C/P paradigm and in the use of planning rhetoric concurrent with a high portion of decisions being made on a disjointed, incremental, remedial basis. The attempt to close this gap has resulted in evaluation, audition, and centralization of decision-making, along with increased attention to hierarchical accountability. Schmidtlein (1974) points out that the broader assumptions underlying efficiency, accountability, and comprehensive planning are increasingly being questioned, and that a new approach to policy formulation is needed, which takes into account the constraints that affect the legitimacy and effectiveness of decision strategies, and at the same time is sensitive to inevitable value trade-offs.

**Formative and summative evaluation** Scriven (1967), observing what he calls a confusion between goals and roles of evaluation among practitioners, developed two new concepts, formative and summative evaluation.

The goal of evaluation is the same in all instances, to obtain some credible judgment related to the merits of a program. Evaluation is a methodological activity, consisting of gathering and combining performance data with a weighted set of goal scales to yield either comparative or numerical ratings, and in the justification of the data-gathering instruments, the weightings, and the selection of goals (Scriven, 1967).

The roles of evaluation, on the other hand, differ. One is to seek judgments for the purpose of improving a program. A diagnosis of the causes for success and/or failure of a program is important for this
purpose. Scriven calls this formative evaluation. The upgrading of the program becomes possible during the ongoing state.

To determine the ultimate worth of a program is the role of summative evaluation. This kind of evaluation is done when a program has reached stability. It pinpoints the merits of a program at a given point of time and can be done at any time. Because of their different roles, formative and summative evaluation necessitate quite different types of data-gathering and analysis strategies.

Scriven asserts that a proper evaluation of a program requires a combination of formative and summative evaluation, as neither by itself yields a full appraisal of the program. He further stresses the role of judgment without which there is no evaluation. The evaluator is best qualified to pass the judgments, especially an external evaluator.

Goal-free evaluation When Scriven formulated his concepts of formative and summative evaluation, he stressed that evaluators must appraise goals. Later, because of concern with side effects of programs, he reversed his position to advocate goal-free evaluation.

Evaluation should concern itself with determining the over-all effects of a program, whether intended or not (Scriven, 1972). In this context, evaluation of stated goals becomes unnecessary and could even be detrimental if there were a tendency (on the part of the evaluator) to look mainly in the direction of acknowledged goals.

The difficulties of evaluating goals stem from several sources. Stated goals, or objectives of a program, are often very different from real goals. Goals are frequently stated so vaguely as to cover both desirable and undesirable activities by almost everyone's standards.
Furthermore, goals are only a subset of anticipated effects; they are the ones of special importance, distinctive to the project. A shift of goals midway in a project renders goal-oriented evaluation meaningless.

Scriven does not say that program goals are unimportant. Planning and production require goals and the formulation of these in testable terms is absolutely necessary for the manager as well as the internal evaluator who keeps the manager informed (Scriven, 1972).

Goal-free evaluation is appropriate for summative evaluation, but as a good formative evaluation gives a preview of summative evaluation, goal-free evaluation is legitimate for formative evaluation as well.

Scriven (1972) considers goal-free evaluation best done by external evaluators, who will be more objective than internal ones, as they are less likely to have a vested interest in the outcome of the program, and less prone to look not only for stated goals but for actual effects. One of the values of goal-free evaluation lies in noticing something that previously has been overlooked, or in producing a novel overall perspective.

Goal-free evaluation has methodological and practical issues. Standards of merit, or criteria, are needed for judging in evaluation. Scriven suggests that needs, properly identified through need assessment become standards instead of goals. Since need assessment is a comparison of the real with the ideal needs are goal-related as far as the ideal is a representation of prior goal statement (Stuffelbeam, 1972). Alkin (1972) points out that goal-free evaluation is not really goal-free at all, but has broader context goals, directed at a different and usually wider
decision audience. He considers the method in essence to be a retro-
spective and nonexplicit need assessment.

The question of gathering data is a problem in goal-free evaluation. It is extremely difficult to observe and collect data about everything. Scriven suggests the use of two external evaluators, working independently until the end of the evaluation. From a practical point of view, this would prove to be a very costly solution, especially as Scriven implies that an internal evaluator is present as well. If the evaluation has to have a focus, will it be according to the evaluator's or the program planner's bias?

In the controversy around the concept of goal-free evaluation, fear has been expressed that it will cause educators to totally abandon goal-setting. Scriven considers goals important in the planning context. Goal-free evaluation might remind educators to look for important results of instruction outside the stated goals. It is a mode of evaluation that might be considered in an innovative program, where goals are vaguely or grandiously stated and apt to change during the implementation. It might also find a place in evaluation of broad-aimed social action programs.

**Stake's congruence model of evaluation** Based on the earlier work of Scriven, Stake developed his congruence model in the late 1960's.

Stake (1967) defines evaluation as the process of describing and judging an educational program. The description and judgment need to be based on formal inquiry. Information data are required about antecedents, transaction, and outcome. These classifications are both relative and absolute; at a given time any item may be either an antecedent, a trans-action, or an outcome. Antecedent is any condition existing prior to
teaching and learning which may relate to outcomes. Transactions are the succession of engagements which comprise the process of education. They are dynamic, while antecedents and outcomes are relatively static. Outcomes are the consequences of educating: immediate and long-range, cognitive and conative, personal and community-wide. Using systems theory concepts, antecedents equals input, transaction equals process, and outcome equals outputs.

Antecedents, transactions, and outcomes need to be both described and judged. Judgmental statements are classified either as general standards of quality or as judgments specific to a given program. Descriptive data are classified as intents (planned-for goals, not in student behavioral terms) and observations.

Every program has its rationale, often only implicit. It needs to be stated and considered separately in evaluation as it provides a basis for evaluating intents, as well as for choosing reference groups.

Figure 1 shows a layout of statements and data to be collected by the evaluator of an educational program, i.e., rationale, intents, observations, standards and judgments about antecedents, transactions, and outcomes.

Descriptive evaluation data can be processed in two principal ways: finding the contingencies among antecedents, transactions, and outcomes; and finding congruence between intents and observation (Figure 2). Data are congruent if what was intended actually happened. Congruence does not indicate that outcomes are reliable or valid, but that what was intended did occur. The relationships or contingencies among variables need to be
Figure 1. A layout of statements and data to be collected by the evaluator or an educational program (Stake, 1967, p. 529)
Figure 2. A representation of the processing of descriptive data (Stake, 1967, p. 533)
studied in order to improve education, especially outcomes that are contingent upon particular antecedent conditions and/or instructional transactions.

Intents are evaluated using logic as the contingency criterion. Evaluation of observation contingencies depends on empirical evidence.

The contingencies and congruences identified by evaluators are subject to judgment by experts and participants (reference groups). The importance of noncongruence will vary with different viewpoints. Perception of importance of congruence and contingency requires the evaluator's careful attention.

The methodology of processing judgments in inadequate. Standards, or benchmarks of performance having widespread reference value, are not in common use. There are two bases for judging the characteristics of a program: with respect to absolute standards as reflected by personal judgments; and with regard to relative standards as reflected by characteristics of alternate programs (Figure 3).

Absolute standards indicate acceptable and meritorious levels for antecedents, transactions, and outcomes. They can be generated from words of experts of standard-setting bodies like professional agencies. The judging act consists of assigning a weight, an importance, to each set of the produced standards.

Rational judgment in educational evaluation is a decision as to how much to pay attention to the standards of each reference group (point of view) in deciding whether or not to take some administrative action.
Figure 3. A representation of the process of judging the merit of an educational program (Stake, 1967, p. 537)
Relative comparison is similar to absolute comparison, but standards are taken from descriptions of other programs. The evaluator selects which characteristics to attend to and which reference programs to compare to. An overall, or composite, rating of merit is possible if both relative and absolute judgments are made (Stake, 1967).

Stake's model is broader in scope than Tyler's model. Process is clearly dealt with, but the procedures for selecting antecedents and learning experiences are less clear. Values are taken into consideration in the analysis of the philosophy of the program, or the rationale. The model also provides for the generation of standards by which the value judgments can be made. The congruency model does not provide for adequate feedback in program development because it focuses on the terminal part of the program. Nor does it provide for evaluation of decision alternatives during the transaction stage (Shiplett, 1974). Stake (1973) deals with these questions in what he calls responsive evaluation.

Responsive evaluation

Many innovative educational projects change on a weekly or monthly basis, and what may at one time be a primary set of goals or objectives will be dropped and replaced by another set. In such programs, Stake (1973) deems that the best an evaluator can attempt to do is to serve as a sounding board and provide quick personal judgment in a form called responsive evaluation.

An educational evaluation is responsive if it orients more directly to program activities than to program intents, if it responds to audience requirements for information, and if the different value-perspectives of
the people involved are referred to in reporting the successes and failures of the program.

Responsive evaluation does not presume that only measurable outcomes testify to the worth of the program. It requires planning and structure, but relies little on formal statements like objective, test scores, and the like. These are not the bases for the evaluation plan but a component of the instructional plan and treated as such. The proper amount of structure for responsive evaluation depends on the program and persons involved. Choice of data-gathering instruments is made as a result of observing the program in action and of discovering the purposes important to the various groups having an interest in the program.

Stake prefers issues to objectives or hypotheses as bases for the evaluation plan, on the ground that issues better reflect a sense of complexity, immediacy, and valuing than do objectives.

Methodology in responsive evaluation consists of continuous observation and feedback to the following recurring events: talking with clients, program staff, and audiences; identifying program scope; obtaining an overview of program activities; discovering purposes and concerns; conceptualizing issues and problems; identifying data needs with regard to issues; selecting observers, judges, instruments, if any; observing designated antecedents, transactions, and outcomes; organizing material according to themes, preparing portrayals and case studies; screening material and matching issues to audiences; preparing informal reports for audience use; and assembling formal reports, if any.
Although responsive evaluation sacrifices some precision in instruments in order to increase utility of findings to persons in and around a program, Stake claims it has purpose and criteria, substantive and functional structure, portrayal and holistic communication, utility, and legitimacy.

Because responsive evaluation is tailored to each particular program, the evaluation results cannot be generalized beyond the program itself, i.e., findings from one program can neither be compared with nor added to those of other programs.

The context, input, process and product (CIPP) evaluation model
The CIPP model was developed by Stuffelbeam, Guba, and Clark. It has its base in decision-making and is conceptualized in terms of a system. Its main audience is administrators at different levels.

The rationale for the CIPP model is that the quality of a program depends upon the quality of decisions in and about the program. The quality of decisions depends upon decision-makers' abilities to identify the alternatives which comprise decision situations, and to make sound judgments of these alternatives. Making sound judgments requires timely access to valid and reliable information pertaining to alternatives. The availability of such information requires systematic means to provide it, and the processes necessary for providing this information for decision-making collectively comprise the concept of evaluation (Stuffelbeam, 1968).

The purpose of the CIPP model is to reduce uncertainty regarding decision choices. Evaluation is seen as having a service function: to
provide the decision-maker with relevant information, thereby optimizing improvements in education. In this light, evaluation is defined as: "the process of delineating, obtaining, and providing useful information for judging decision alternatives" (PDK Committee, 1971, p. XXV).

**Basic components** Basic components of the CIPP model are the decision settings with the most relevant decision strategy for each setting, the four types of decisions, and the corresponding four types of evaluation.

The CIPP evaluation model makes an adapted use of the decision-making methodology in political science, as conceptualized by Braybrooke and Lindblom (PDK Committee, 1971). A matrix can be formed by crossing two continua; small versus large educational change, and high versus low information grasp to support change. Thereby, four different decision-making settings, each with different activity, purpose and basis are identified (Figure 4).

The metamorphic decision setting is primarily of theoretical value, as a complete change, based on total understanding is utopian in nature. It is not dealt with in the CIPP model, although the PDK Committee (1971) points out that many educators act as if they worked within this setting.

The homeostatic decision setting, denoting high information grasp and small degree of change, aims at maintaining the normal balance in an educational system, and is based on technical standards and routine, cyclical data collection. This is the most prevalent decision setting in education.
Figure 4. Decision-making settings (PDK Committee, 1971, p. 62)
Most relevant to the homeostatic setting is the synoptic ideal decision model, aspiring to a high degree of comprehension. It requires the consideration of all possible consequences for all possible alternatives in terms of all relevant criteria. The decision-maker then chooses the best alternative from all possible alternatives. In reality, it is not possible to specify with certainty all alternatives and criteria or to predict with certainty or with known levels of risk all the possible consequences associated with each alternative in terms of all the relevant criteria.

The incremental decision setting, with low information grasp and low degree of change, includes developmental activities for the purpose of continuous improvement. It makes use of expert judgment and structured inquiry. Many so-called innovations, attempting to make improvements without risking a major failure, fall in this category. The disjointed, incremental decision model is relevant to this setting. The method used is problem analysis, a tentative selection of apparent change possibilities and through a trial of these, a successive approximation of a solution is reached. The use of this model affords continuous efforts to improve a program without the risk of major failure.

The neomobilistic decision setting occurs when a large degree of change is aimed at, but there is little or no support of theory and knowledge. This setting is present in inventing, testing and diffusing new solutions to significant problems. The neomobilistic change effort characteristically begins as an exploration. Risk-taking is generally high.
The planned change model, proposed by Clark and Guba and extended by Stuffelbeam (PDK Committee, 1971), serves the purpose of large change. Many steps and agencies are involved over a relatively long timespan. The steps are research, development, diffusion, and adoption, each of which is a process in itself. When used, a specific change procedure must be designed in detail, and the planning done well in advance. For success, there is need for a large amount of central control and coordination.

Although many different decision models exist, each decision setting has one that has special relevance, as specified above. A decision-maker is likely to choose a decision model unconsciously, guided by his own behavior patterns and the decision process setting. Yet, much of the dysfunction which occurs in educational change is believed to be due to the choice of an inappropriate decision model for a particular decision setting.

The PDK Committee (1971) identified four major types of decisions to be made in an educational program or project: (1) planning decisions, specifying major changes that are needed in a program; (2) structuring decisions, detailing means to achieve the ends established as a result of the planning decisions; (3) implementing decisions, or those involved in carrying through the action plan; and (4) recycling decisions, determining the relationship of attainments to objectives and whether to continue, terminate, evolve, or drastically modify the activity (Figure 5).

Corresponding to these four types of decisions are four types of evaluation. The first is context evaluation, with the purpose of providing a rationale for determining objectives. It defines the relevant
<table>
<thead>
<tr>
<th>INTENDED</th>
<th>ACTUAL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ENDS</strong></td>
<td><strong>MEANS</strong></td>
</tr>
<tr>
<td>PLANNING DECISIONS</td>
<td>STRUCTURING DECISIONS</td>
</tr>
<tr>
<td>to determine objectives</td>
<td>to design procedures</td>
</tr>
<tr>
<td>RECYCLING DECISIONS</td>
<td>IMPLEMENTING DECISIONS</td>
</tr>
<tr>
<td>to judge and react to attainments</td>
<td>to utilize, control and refine procedures</td>
</tr>
</tbody>
</table>

Figure 5. Types of decisions (PDK Committee, 1971, p. 80)
environment and describes the desired and actual conditions pertaining to that environment, identifies unmet needs and unused opportunities, diagnoses the problems that prevent needs from being met and opportunities from being used. The second type is input evaluation, with the purpose of providing information needed for determining the utilization of resources to meet program goals. It is carried out through the identification and assessment of relevant capabilities of the responsible agency, of strategies for achieving goals, and of designs for implementing a selected strategy. Decisions about program design are based upon input evaluation. The third type is process evaluation, which aims at detecting or predicting defects in the procedural design or its implementation during the implementation stages; at providing information for programmed decisions; and at maintaining a record of the procedure as it occurs. The last type is product evaluation, with the purpose of measuring and interpreting attainments, at the end as well as during the project cycle. Traditionally, evaluation has focused on product, but only at the terminal end of the cycle. In this model, product evaluation is also done at intermediate levels in order to assess contributions to the achievement of ultimate goals. Product evaluation helps the decision-maker when making recycling decisions.

The CIPP evaluation model provides the conceptual basis for applying the theory of evaluation developed by Stufflebeam, Guba, and Clark (Figure 6).

The rectangles represent the four kinds of evaluation. The small loops attached to these stand for the process of delineating, obtaining, and providing information.
Figure 6. An evaluation model (PDK Committee, 1971, p. 236)
The hexangles represent the four types of decisions served by evaluation. The large E represents the decision settings.

The outer cycle represents a continuous, systematic context evaluation mechanism that provides both congruence and contingency date. If context evaluation indicates no discrepancies between the intention and actualities, or between possibilities and probabilities, the planning decision might be to continue at the same level. If not, a rational decision-making body might plan for change, either neomobilistic, incremental or homeostatic.

The homeostatic change leads straight to structuring decisions, installation, adjustment of the context evaluation mechanism, and to program operations.

Neomobilistic and incremental change lead to input evaluation, from which it generally goes to structuring decisions, and thereafter to trial. However, it is possible to bypass the trial stage and go directly to the installation of change procedure. The trial goes through both a process and a product evaluation. The process evaluation leads to an implementation decision, which might be to renew the trial, or to a recycling decision. The product evaluation leads to a recycling decision, which in turn can take many paths. It might lead to an implementation or to a termination decision. It might also retract to a new trial, or to a new structuring decision, to a new input evaluation, or to a new planning decision.

Among the features of the model pointed out by the PDK Committee (1971), it is noteworthy that it considers that context evaluation should
be conducted in a general undifferentiated setting, while input, process and product evaluation occur within specific, suboptimized settings; that leadership personnel in the context evaluation mechanism should have a continuing and direct relationship with those who formulate policy and planning decisions; that input, process and product evaluation studies should be more comprehensive and highly structured for neomobilistic change than for incremental change; that input, process, and product evaluation to support homeostatic change generally do not require formal evaluation studies by evaluation units; that process and product evaluation should be carried on simultaneously, with process evaluation receiving heavier emphasis early in a trial and product evaluation receiving heavier emphasis as the trial progresses. The PDK Committee (1971) thinks that evaluation and decision-making functions should be separate. To insure the objectivity of both the decisions and the evaluation, it is important that information be collected, organized, analyzed, and reported by persons trained in these functions, but they should not be responsible for the strengths and weaknesses of the program.

The PDK Committee (1971), in contrast to thoughts on the general decision-making process, includes a fourth step, action, in the educational decision-making process. Its reasoning is based on the notion that evaluation should service decision problems, i.e., decision-making in problem solving settings.

Two kinds of criteria, consequential and instrumental, are discussed by the PDK Committee (1971). Consequential criteria are defined as: "criteria pertaining to the fundamental conditions being sought" (p. 353),
and instrumental as: "criteria related to the accomplishments that are intermediate level and that contribute to the achievement of ultimate objectives" (p. 354).

**Methodology** The PDK Committee (1971) rejects the notion that the methodology of evaluation is identical to the one of research, on the grounds that their purposes are different. The purpose of research is to provide new knowledge, universally valid, regardless of its utility, whereas the purpose of evaluation is to delineate, obtain and provide information for making educational decisions, valid and useful within the decision-making context. Research provides certain kinds of information under certain conditions. Evaluation is concerned with responding to a wide variety of needs for information by educational decision-makers under diverse and often adverse conditions.

The focus of evaluation in the CIPP methodology is on different administrative levels. The PDK Committee (1971) distinguishes between two basic levels of information needs for decision-making: the microscopic level spotlighting specific elements in a total system, e.g., the individual student, classroom, or school; and the macroscopic level, concentrating on overall appraisal of the total system, e.g., the district, state, or nation. Each level, needing qualitative as well as quantitative information, requires divergent evaluation techniques and criteria. The purpose of micro level evaluation is to discover system disturbances, or process control. It seeks to aid the decision-maker to increase congruence between intended and actual performance within the system. Its concern is with the incremental decision setting. Macro level evaluation
is more future-planning oriented, seeking to redesign or totally replace the existing system. This type of evaluation occupies itself with innovations, and its decision setting is neomobilistic.

The CIPP methodology is centralized on the process of delineating, obtaining, and providing useful information for judging decision alternatives. These actions consist of logical units of activity, each with a series of tasks. To understand the methodology, the main concepts in the definition above, as seen by the PDK Committee (1971, pp. 353-355), are outlined below.

Information is the aggregation and arrangement of data elements so as to reduce the uncertainty on the part of the decision-maker. Data, representations of observations of an event, become information only when so presented as to reduce the uncertainty for the decision-maker.

Delineating consists of identifying evaluation information required through an inventory of the decision alternatives to be weighted and the criteria to be applied in weighting them.

Obtaining means making information available through such processes as collecting, organizing and analyzing, and through such formal means as measurement, data processing, and statistical analysis.

Providing is fitting information together into systems or subsystems that best serve the purposes of the evaluation, and reporting the information to the decision-maker.

Useful denotes satisfying the scientific (internal and external validity, reliability, objectivity), practical (relevance, importance, scope, credibility, timeliness, pervasiveness), and prudential
(efficiency) criteria, and pertaining to the judgmental criteria to be employed in choosing among the decision alternatives.

Judging signifies the act of choosing among the several decision alternatives; the act of decision-making.

Delineating and providing information is largely a matter of interaction between the evaluator and the decision-maker; the obtaining of information is the task of the evaluator.

The evaluation methodology is the methodology of an information system designed to provide information for project, program, and system decisions. The design, proposed by the PDK Committee (1971, pp. 156-157, p. 213) is based on a work breakdown of evaluation activities. All evaluation studies, context, input, process, and product, should follow the same general course, namely:

1. Delineation of information needs
   1.1. Definition of system
   1.2. Specification of decisions
   1.3. Statement of evaluation policies
   1.4. Statement of evaluation assumptions

2. Plan for obtaining information
   2.1. Collecting data
   2.2. Organization of data
   2.3. Analysis of data

3. Plan for providing information
   3.1. Preparation of reports
   3.2. Dissemination of information

4. Evaluation of the evaluation

Each sub-task can be further broken down, e.g.,
1. Delineation of information needs
   1.1. Definition of system
      1.1.1. Model of the system
      1.1.2. Elements of the system
      1.1.3. Characteristics of system elements

It is important in the CIPP model that provided information is of the
same level as the decision. Ward and Lehman (1972) suggest a process for
describing levels in a decision setting. Information in use, the flow of
that information within the decision structure, and the decision structure
itself need to be outlined. The authors base their analysis and procedure
to a great extent on statements from program personnel, in which they de-
fine their roles in specific decision functions as well as in cooperation
with whom. This method might delineate the formal decision structure, but
may or may not tap the informal decision structure and information flow.

Stuffelbeam (1971) claims that the CIPP model, although originally
developed to provide timely information in a systematic way for decision
making, can also serve the retroactive purpose of providing information
for accountability. He defines accountability as, "the ability to account
for past actions in relationship to the decisions which precipitated the
actions, the wisdom of those decisions, the extent to which they were
adequately and efficiently implemented, and the value of their effects"
(p. 20). Administrators are thus responsible for both ends and means of
their programs. The CIPP model can produce data for accountability pur-
poses through the records kept of objectives and bases for their choice;
of strategies and design chosen and the reasons for their choices; of the
actual process, as well as of the attainments and recycling decisions.
Both internal and external evaluation is required in the CIPP model (Stuffelbeam, 1971). The internal evaluation systematically supports the decision-making and provides the basic data for accountability. External evaluation periodically brings in outsiders to ask important questions and make objective assessment, providing independent, summative evaluation of the system's goals, designs, procedures, and results. The internal evaluator provides much data needed by the external evaluator, who in turn can enhance the credibility of the internal evaluator.

**Advantages and disadvantages of the CIPP model**

Randall (1969) discusses difficulties in applying the decision-making evaluation theories, which are based on the assumption that the most effective decisions are those based on the best information. The key to effective operation of a system based on the CIPP model for evaluation is that decisions are identified along with the persons who are involved in the decision process. Then a systematic attempt is made to get relevant, timely information in a usable form to the decision-maker. Decisions to be made are not always easily recognized, not even by the decision-makers themselves. Decision criteria may change as new developments occur, as new information is obtained, and as conditions change. The passing of time may also cause constraints to change. Since there is always some lag between the time when decisions are identified and the time when information is collected, processed, and reported, the system must continually be alert for changes in constraints that might change the basis on which decisions will be made. Key persons involved in any strategic decision need to be identified and supplied with information. There are problems
with the timing of information, to get the best information possible in the time allotted. Hardship may be encountered in identifying what is relevant information. Final criteria for this are what the decision maker considers relevant, or it will have little, if any, effect on a decision. It might become a Herculean task to report information in a useful form, to select the right degree of sophistication and specification. Criterion will be the length of time the decision-maker will be likely to have to consider the information, as well as his competence in understanding terminology and techniques used. Some information might have to be presented in several different forms to different decision-makers. Furthermore, the communication and interaction between key decision makers and evaluator is essential in decision making evaluation, the establishment and maintenance of which has many obstacles to overcome in most organizations.

Lange (1974) argues that the process of delineating, obtaining and providing useful information for judging decision alternatives cannot be said to be evaluation at all. Evaluation involves judgement, hence the evaluation takes place in the judgment process when the data are analyzed, synthesized and mixed with personal values to produce judgment.

He further points out, that to include context assessment, or need assessment, within the concept of evaluation creates an overlap between planning and evaluation that on one hand is apt to create confusion on the part of the practitioner, who needs to know if he is collecting data for evaluating the school or for supporting a specific set of planning decisions. On the other hand, it has alerted a wider segment of educators
to the desirability of generating a data base for identifying needs and selecting objectives for improvements.

The idea of context evaluation is useful, but it provides no new techniques for gathering data. It relies on those strategies well tried and used for some time by researchers and/or survey specialists. Added to that, Lange (1974) maintains that no educational agency has been successful in establishing and maintaining a general, wide-ranging assessment mechanism that covers every base that ought to be covered by the agency.

Lange (1974) separates out as good features of the CIPP framework the distinction made between process and product evaluation, and the concept that the two be performed simultaneously. Another distinction is the implication that there are a variety of change settings within which an agency might find itself, and that not all of the change settings, i.e., those with small changes and where the result of the change is well understood and can be readily predicted, require detailed evaluation.

Applications and adaptations of the model In spite of criticism, the CIPP model has been used in many contexts and for a variety of purposes; in its original form, or modified.

Shiplett (1974) adapted the CIPP model of evaluation in an attempt to increase its effectiveness when used in adult education. He chose to work with the CIPP model because he deemed it provided the most holistic approach to evaluation; served both in planning and in implementation of an educational program; provided for evaluation of process, including group-process, with continuous feedback and of procedures both in input and process evaluation; and because it did not neglect the value question. Shiplett (1974) felt a need for adaptation because in the CIPP model, the
decision process is viewed in an essentially nonpolitical, rational way. In adult education, decision-making is considered a social process, which needs a conflict management strategy rather than the centralized, rational one. Furthermore, the CIPP model depends on decision rules, whereas adult education decision-making, according to Shiplett (1974), relies on what he calls ameliorative themes.

The adaptation was done in six steps. The first included developing criteria for evaluation in adult education through a survey of adult educational literature and an examination of the compliance structure in the normative, cultural setting according to Etzioni's typology of organizations (cited in Shiplett, 1974). The latter is a type of organizational setting which serves cultural goals. It is characterized by moral involvement of the participant, based on personal values and expressed in personal commitment. Compliance is based on power and involvement. The normative power rests upon the allocation, manipulation, and deprivation of symbolic rewards. Cultural goals aim toward general well-being, i.e., self-actualization, community improvement.

The second step consisted of identifying features of an effective evaluation model of adult education through the application of the developed criteria. Shiplett (1974) identifies six needs of adult educational evaluation: to allow for a high degree of self-evaluation, for active involvement of the participants, and for participant evaluation of the group process with the possibility of immediate feedback. Others are to allow for process as well as product evaluation, for active involvement
of interested participants in the evaluation during planning, and in the areas of administration and program development.

Analyzing the CIPP model in terms of its understanding of educational change, and educational change as understood in adult education became the third step. Thereafter followed the changing or elimination of features of the CIPP model that were not compatible with the understanding of change in adult education. Shiplett (1974) argues that all planned change in adult education is incremental in nature because adult education is situational and the participants involved in the designing, planning and evaluating of their own program. The strategy needed is the disjointed, incremental one. He rejects the neomobilistic setting with its strategy of planned change on grounds that it is unlikely that adults will voluntarily adopt this strategy, and that the emphasis on local program development in adult education almost excludes the development and testing of pilot projects and the diffusion of these to new situations. The dynamic development of a program occurs after it is installed, according to Shiplett (1974). Therefore the E-shaped figure in the CIPP model is omitted and with it the whole structure designed for pilot tests. Figure 7 shows the adapted model.

Shiplett (1974) recognizes three domains of concern operating in the planning and implementing of an adult educational program: the administrative, the developmental, and the participant.

A clear distinction is also made between two phases of educational programs, the planning stage and the implementation stage, whereas the CIPP model deals with the trial and the installation phases. Each stage
Figure 7. The CIPP evaluation model adapted to adult education (Shiplett, 1974, p. 129)
has its own appropriate kinds of evaluation, and its own types of decisions to be made. The planning phase consists of context and input evaluations to serve planning and structuring decisions respectively. Its tasks are in both the administrative and developmental domains, with objectives in each. The implementing phase comprises the process and product evaluation to serve implementing and recycling decisions. Tasks for the implementing phase exist in all three domains, administrative, developmental, and participant.

Feedback becomes crucial with three different domains of concern operating in each stage and when decision-making uses conflict management strategy. It is needed between, as well as within, the different domains and stages.

The adapted model, assuming that all planned change in adult education is of incremental nature, tends to have only on-going or recurring educational programs with voluntary participants in mind. Once installed, the program development is dynamic. The initial establishment of a program is less adequately dealt with, apart from need assessment. Innovations aiming at larger changes, i.e., at a completely different audience, a completely new area of content, do occur also at a local level. Nor is all local adult education independent of the larger social system; state and federal influence is a part of the local decision-making, both at the administrative and developmental domains of concern.

Another application of the CIPP model was made by Reeves and Michael (1973) when evaluating a dental team training program involving expanded use of paraprofessionals.
Context evaluation, in the form of need assessment through interviews and literature review, formed the basis for making planning decisions, i.e., helped to establish what the goals and objectives for the program should be, and led to the actual formulation of seven preliminary goals, expressed in student behavioral terms.

Input evaluation provided information regarding such variables as the legality of the approach, existing programs, job descriptions of auxiliaries and dental hygienists, available resources and anticipated constraints on these, and results of pilot studies. Based on this information, structuring decisions were made concerning the design of the program; a synthesized job description, allocation of human and physical resources within the constraints, student selection, plans for both process and product evaluation, and instruments to be used in the evaluation.

The process evaluation led to decisions regarding alternative courses of action that could be taken to maintain and improve the quality of the program. Evaluators obtained information about strengths and weaknesses of participants in both technical and human relations and management skills, manifested at various points of time through observation scales, and about congruence between desired and actual attainments on a daily and/or weekly basis. The extent to which facilities, instructional materials and supplies were used in prescribed ways and amounts were assessed as was the degree to which roles of participants were understood and implemented. The degree to which misunderstandings about job functions and responsibilities prevented the fulfillment of program objectives, as well as the patients' attitude towards receiving experimental
dental services were also investigated. Implementing decisions made included adjustment in specification of task requirements, modification of existing and design of new evaluation instruments, and increased emphasis on the development of human relations and management skills.

The product evaluation was made nearly simultaneously with the process evaluation. At the terminal points of training sessions of various teams, information about the capabilities of the team and team members was obtained through performance criterion scales and various observation scales. Cost and effectiveness data were procured.

The paper (Reeves & Michael, 1973) is a description of the application of the CIPP model of evaluation. The program is called innovative, but mentions results of pilot studies as part of input evaluation. The nature of these are not disclosed. No information as to the completeness, reliability and validity of information gathered at various stages is provided, and none about the evaluator(s), i.e., external, internal, degree of involvement in program, degree of communication evaluator/decision-maker. Therefore, it is difficult to judge the quality of the evaluation described. On the other hand, the CIPP model is a framework and does not insure quality per se, but aims at reducing uncertainty on the part of the decision-maker. The above evaluation did produce a variety of data.

Moburg (1973) designed a study to demonstrate the application of the CIPP model as a means of making decisions related to the competence of teachers in determining student functional reading levels. It was used to identify decision points and appropriate procedures in the investigation of the problem. Context evaluation was performed to answer the question:
Can selected elementary classroom teachers determine a student's instructional and independent reading levels? Information was obtained through literature review, a needs assessment questionnaire, and a criterion instrument. Input evaluation was designed to reduce uncertainty in making decisions in regards to the question: What is the most viable alternative in meeting the need as identified? Process evaluation responded to the question: What procedural defects or potential sources of failure are present in the training program and how can these problems be met? As a result, a training package, aimed at providing training in marking, scoring and interpreting oral reading errors was developed. Product evaluation inquired about the effect the experimental program had on the ability of teachers to determine student functional reading levels. A post-test-only control-group design was used to produce product evaluation data.

Moburg concluded that the CIPP was a useful, effective approach in dealing with a specific problem of teacher competence.

Several studies have used only parts of the CIPP theoretical framework. Rorie (1973) used Stuffelbeam's decision-definition of evaluation when attempting to devise a simple and objective system for use in making comparative evaluations of similar type vocational and technical education programs offered at different public community colleges. The system consisted of a comparative evaluation model and a comparative rating scale model. Steps included in the models were: identification of decision, of needed information, and of criteria for judging; selection of measures for criteria; weighting of measures by a panel of decision-makers; collection
of data; calculation of program rating scores; analysis of scores and recommendations for revision.

Hunt (1975) applied the CIPP point of view to create a model for the evaluation of federally funded postgraduate programs on family planning as delivered to practicing physicians in five medical schools. The students, curriculum, faculty, and teaching and housing facilities were viewed as a system bounded by a geographic entity. Each school chose its own emphasis and methodology, yielding four different teaching models. The new evaluation model, applicable to all teaching models, was developed and information for context, input, process and product evaluation was delineated and obtained using selected measurement instruments. Criteria for judging the worth of such programs were examined and discussed.

Griessman (1969a) used a modified version of the CIPP model in evaluation of the Concerted Services in Training and Education pilot projects in Arkansas, Minnesota, and New Mexico. The program attempted to stimulate area development through coordination of services and programs at local and national levels. The research model used the following four stages of evaluation: (1) context (environment) - describing the area where changes are to occur, delineating the area's unmet needs, and analyzing the problems that underlie those needs; (2) input (resources) - listing relevant capabilities of appropriate programs as well as strategies that might be manipulated in an action program; (3) process - investigating procedures that are actually being employed; and (4) product - assessing the extent the program's objectives have been attained. In addition, account was taken of unanticipated effects. Procedures used in the evaluation were
selecting an evaluative research model; applying objectives and identifying assumptions; selecting research strategies, collecting data, and analyzing and interpreting data. The above study seems to have used the framework of the CIPP model, but to have changed the methodology as evaluative data were collected after programming decisions were made and implemented.

Kohl and Achilles (1970) extended the CIPP model with elements from two other evaluation models. Their resulting model included four separate phases: (1) context, status or assessment evaluation; (2) input or planning evaluation; (3) process or operation evaluation; and (4) product or final evaluation. Within each of these stages procedural outlines were developed and presented in the form of flowcharts. Each stage of the evaluation design is seen as a logical extension of and a complement to planning activities. Planning establishes the values and the philosophical point of decisions as well as the direction toward which the organization should move. Planning is seen as preparation for decision-making and providing the framework for action. Evaluation serves as a review of decisions in light of certain data, criteria, or changes that can be documented, and as a procedure for exploring the need for new decisions.

Although labeled an extension of the CIPP mode, Kohl and Achilles (1970) differ in at least two main aspects. They clearly separate between planning and evaluation, whereas in the CIPP model these two concepts merge. The role of evaluation in the CIPP model is to delineate, obtain and provide useful information for judging decision alternatives, while
Kohl and Achilles (1970) use evaluation to review decisions and to explore the need for new decisions.

In reviewing studies purporting to use the CIPP model, its framework, i.e., the assessment of context, input, process, and product, seems to be a useful tool in evaluation. There appears to be a trend to separate the planning and evaluation stages, and a tendency on the part of the evaluators to modify the methodology. The latter often take the form of changing the timing of the provision of information in relation to especially process and product decision-making, i.e., to provide information for accountability rather than for decision-making.

Broad-aimed programs and evaluation

Notably since the 1960's, educational and social action programs increasingly fall under the category of broad-aimed programs. Weiss and Rein (1970) define broad-aimed programs as those which hope to achieve nonspecific forms of change-for-the-better, and which also, because of their ambition and magnitude, involve unstandardized, large-scale interventions. State and federal programs evaluated in only a few sites are also included in this category. Farmer (1975) considers educational programs to be broad-aimed if they seek to bring about cognitive, affective, and psychomotor changes that result in changes in actual performance and in the quality of the participants' lives.

As broad-aimed programs have much in common with innovative programs, evaluation of these will be discussed under the latter heading.
Innovative programs and evaluation

Education, as a means to improvement, is trying out innovations. The concept of what constitutes an innovation varies, but generally, innovations do not necessarily mean invention, and the innovation is relative to the individual or locality.

Rogers and Shoemaker (1971) define an innovation as an idea, practice or object perceived as new by an individual. It is a subjective definition, based on the assumption that what matters is whether the idea is new to the receiver, rather than whether it is objectively new in the eyes of experts.

Brack (1975) considers a program innovative if it is being offered to an entirely new audience, or it employs a different methodology, or the context in which the project operates has a new dimension.

The evaluations of innovative programs have many features in common with the evaluation of ongoing and established programs, but because of heightened interest in the results of the study, the more dynamic politics surrounding the project, and the frequent involvement of new staff members, it also has some unique characteristics.

Sometimes innovative programs are implemented before the outcomes are specified or the outcomes might change as the program develops. The evaluation design needs to be able to accommodate this dynamic state of development.

As innovative programs have no forerunners, the question of setting performance level becomes a problem. Not to specify expected performance level until data are available from the program leads to the temptation of
setting the expected level at the actual level, thus thwarting the realization of the full potential of the new program.

All programs have effects other than the intended ones. In established programs positive side-effects are incorporated into the program, and negative ones ignored. A crucial issue for evaluation of new programs is whether to include the stated outcomes only in the appraisal, or to embrace additional effects as well.

New programs are prone to the "Hawthorne effect", or to have effects beyond what would occur under normal conditions. Being involved in an experimental program might reduce boredom, change routine and habits on the part of program personnel and participants, and thereby increase their efforts. Those involved in new programs often try harder in order to prevent the project from failure. Administrators are also apt to give more support to innovations than to well established programs.

Another pitfall in assessment of innovative programs is that innovations, aiming at changing value systems and customs, are subject to what Brack (1975) calls "sleeper effect". Results are delayed and will not show up during early evaluation studies.

Suchman (1970) makes a clear distinction between demonstration and operational programs on the basis that their objectives are different and hence their research designs must be dissimilar.

A demonstration program is usually a one-shot trial in order to either develop a model or prototype for future operational programs, or to test the effectiveness of an operational program on a small scale. The
research design most appropriate for demonstration programs is the experimental one, the only design that permits the drawing of cause-effect conclusions.

An operational program, on the other hand, has already established its worth and focuses on improvement of services. Pure experimental design is less feasible in social action programs. It has to be adapted to each particular situation, but in doing so, the evaluator should endeavor to approximate the experimental model as closely as possible.

Suchman (1970) further differentiates among three types and levels of demonstration programs. One is the pilot program, in which new approaches and new organizational structures or procedures can be tried out on a flexible and easily revisable basis. This program requires a quick-and-easy type of evaluation with primary emphasis upon feedback for program changes. The second is the model program, which is the outcome of a series of pilot projects. This program is designed as a definitive experiment to test the hypothesis that activities A,B,C can achieve objectives X,Y,Z. The research design is a carefully controlled experimental one. When evaluating the model program, conclusions cannot be directly generalized to standard programs; another step in program development must be taken. That is the prototype program, in order to find out if the results of previous testing is practically and realistically feasible on a large scale in terms of available resources. Here the evaluation design will attempt to resemble the experimental approach, comparing the prototype program with traditional programs as controls. The prototype program should be evaluated under conditions as similar as possible to the pro-
posed operational program in order to permit the results to be applicable to these programs.

In an innovative program, the PDK Committee (1971) recommends that the process and product evaluation be carried on simultaneously, with process evaluation receiving heavier emphasis early in a trial, and product evaluation increasingly stressed as the trial progresses. Macy (1975) elaborates on this idea, suggesting that the emphasis changes in accordance with the implementation success which the program achieves. In general, in the beginning of the developmental phase, when a program encounters implementing difficulties, evaluation should give process evaluation high and product evaluation low emphasis. As the program approaches optimal implementation, that is, when program structure is theoretically sound and program personnel function within the specified program structure, process evaluation should have low emphasis and product evaluation high. During the maintenance phase of the program, a continuously low-level emphasis of process evaluation and periodic high-level emphasis of product evaluation is called for. If during the maintenance stage there are sub-areas of program development, the evaluation principles above should be operative within these sub-areas. Process evaluation must be able to identify the degree of optimal implementation attained, e.g., detect unsound program structure as well as disparities between the operating specifications of program structure and the behavior of program personnel. The outcomes of process should determine the extent of product evaluation conducted in the evaluation of new and developing programs, as these can be used to validate the interpretation of product evaluation.
information. Where process evaluation is not included, there is little or no basis for interpreting product evaluation outcomes, and there is little or no information which will enable program management to improve program development, especially at the local level.

Flynn (1972) sees the need for evaluating educational innovations. On the other hand, as the assessment of initial and early innovations frequently provide evidence which is too premature to make critical decisions regarding the innovation, no program should be rejected out of hand because of a poor showing early in its development. Such a program should be allowed to develop out of the mainstream of education, using volunteer students only.

Charters and Jones (1973) attack the method of evaluating innovations by comparing student outcome data in experimental classes with carefully matched control groups only, because results can be completely misleading. The experimental and control programs need to be described and their actual differences measured and certified. The differences between what the research regards as "experimental" and "control" programs might be fictional, as the path between all announced intention to change and demonstrable changes in student proficiency is long and hazardous.

The authors distinguish between four levels of reality of a school's program, ordered according to their proximity to the attributes of students which they are designed to effect. The first level is the Institutional Commitment, as evidenced and described by authoritative pronouncements, statements of intention, promises. The second level, Structural Context, refers to those changes in final arrangements and physical
conditions that form the context within which each staff member carries out an educational program. Usually this level can easily be documented, but it might be difficult to order the features generically. The third level consists of Role Performance as the behavior patterns of teachers must be observed to change, and change so as to accord with the role performance required by the innovation, before it can be said that an innovation is actually implemented. Institutional commitments and structural changes may not only not cause any behavioral changes of teachers, but may also induce behavioral adaptations on the part of staff members that are contrary to intentions of the innovation. The fourth level constitutes the Learning Activities. It is the students' own activities and experiences that are most immediately related to learning outcomes. The best teachers can do is to induce students to engage in activities deemed instrumental to covert psychological processes they hope to affect.

The state of the program at each level is the product of many forces, only a fraction of which are subject to intentional manipulation and rational planning. Plans have a habit of going astray during their implementation, producing side-effects and unintended consequences, some of which may work in opposite direction to the one sought for.

Campbell (1970) supports the use of the quasi-experimental design and the use of multiple measures, each recognized as having its own components of systematic bias and irrelevancy in addition to random errors, when appraising social innovations, as well as multiple modes of measures with multiple aspects and perspectives. The experimental method, measuring
only one or a few variables can distort the result, as can the treatment of those variables as perfectly and exhaustively relevant.

The issue in the evaluation of broad-aimed programs is not "Does it work?" but "What happened?" (Weiss & Rein, 1970). Evaluation of broad-aimed programs ought to describe the form the program takes, and describe and identify the forces which shaped the program, the nature of the opposition encountered, the reasons for success or failure, and the program's unanticipated consequences. Experimental design creates technical and administrative problems in the evaluation of broad-aimed programs because of its inability to handle critical issues. Among those issues are the difficulty in selecting satisfactory criteria, dealing with essentially uncontrolled situations and with nonstandardized treatments, and the need for information on a variety of variables. Nevertheless, Weiss and Rein (1970) found the experimental design justified in this type of program if one aim, or a single objection, is important enough to justify collecting data which will lead to a relatively unquestionable conclusion.

Issues and problems in program evaluation

Issues and problems in educational and social action program evaluation seem to form two overlapping clusters. The first one deals mainly with administrative questions such as purpose and utilization of evaluation. The second cluster contains technical questions, or methodological problems and those connected with the generation of satisfactory data.

Definition and purpose of evaluation

When surveying literature on program evaluation a reader becomes increasingly aware that authors define and perceive the purpose of evaluation differently. Schools of thought
can be identified, but there seems to be no commonly agreed upon concept of evaluation.

Although a few writers maintain that evaluation consists of factual description only, most authors agree that evaluation involves judgment. The dissension evolves around questions like by whom? at what point? to what extent? using what criteria?

The purpose of evaluation is viewed as diversely as the definition from which it is derived. Because the purpose of evaluation is varied, evaluation and the type of evaluative data collected need to change accordingly. Hence the call for different types of evaluation.

Evaluation is a rational enterprise, often undertaken for nonrational, or at least noninformational reasons (Weiss, 1972a). Guba (1975) distinguishes between overt and covert foci of evaluation. The overt foci, easily found in the literature, are objectives, effects and decisions. Covert foci are likely to remain hidden even under intensive inspection, but may carry more importance than the overt ones. Examples of covert foci are compliance with required evaluation and ratification of a decision already made.

**Evaluation and administrative levels**

Expectations for the evaluation generally vary with a person's position in the system. Top policymakers need information on overall effectiveness of the program, and directors of the program are concerned with the efficiency and economy of general strategies as well. Direct-service staff, dealing with individuals and/or small groups, have practical day-to-day concerns about techniques (Weiss, 1972a).
Cohen (1970) points out that to assess program at macro-level using micro-level data, i.e., in terms of student behavior changes, is a futile undertaking. Social action programs have a broader aim, that is, to bring about political and social changes. A measure of student achievement might be a valid measure of program success only if it stands for other things, i.e., if more knowledge leads to more years of schooling, to better job opportunities, to more money, to more social and economic status. But, as he indicates, there is no conclusive information about the causal relationship between school children's achievement and their later social and economic status, once inherited status is controlled for.

The PDK Committee (1971) identifies five major problem areas that occur with the shift in focus on administrative levels. The first is census versus sampling data, the second centers around specificity versus over-all information. A third area of problems is in the conflict of purposes, the fourth is caused by the confusion of congruence versus contingency evaluation. The final problem area is concerned with baseline data, the purpose of which at the micro-level is to permit judgments about program effects, at the macro-level to make predictions for the future.

**Evaluation and politics** Evaluation is now becoming increasingly political. Policy decisions sometimes hinge on whether evaluation shows good results from an action program or not (Weiss, 1972a). Evaluation is designed to yield conclusions about the worth of programs, and to contribute to decision-making and thereby influence allocations. Therefore, it has the potential to change power relationships and can be viewed as a
political activity. The bigger the program, the greater the likelihood for overt appearance of political competition (Cohen, 1970).

Policy formation originates with the values of the policy-maker and society, and these values are most immune to feedback (Worthman, 1975). Values, through policies, seem to influence the research to be evaluated, but the result of such research appears to have little impact on these values. Cook and Lange (1975) agree that for multiple reasons, evaluation information is seldom utilized for making adjustments to existing policy, let alone for serving future policies. Once a public stand is taken and a policy becomes associated with an individual or group of individuals, the issue tends to become closed and the policy is seen as set and relatively permanent. Furthermore, because of continuously changing context and value orientation in the social setting, evaluation results may arrive too late to be of value for the policy maker.

The American political system is pluralistic, designed to accommodate different values and power centers. Hence multiple goals are often built into federal programs as a result of compromises and trade-offs between competing interest groups (Bernard, 1975). Educational evaluators often confound narrow program questions with broad public policy issues, or fail to recognize these issues. Berlak (1970) suggests tentative criteria for identifying the latter.

Traditionally decision-makers have not accorded evaluative research a major role in policy formation and change in social programming. Informal, rather than formal evaluation has been employed. Administrators may consider evaluative research expensive and of little practical value,
as well as having covert reasons for resisting formal evaluation (Caro, 1971). Ambiguous results help to protect administrators where there is a possibility of failure.

Even the most carefully designed and executed evaluation research does not automatically lead to meaningful action. Disregard for results of evaluation stems from various sources (Caro, 1971). There are limitations of the research itself, a dilemma of rigor, timing and utility. Results can often not be produced early enough to be a major factor in short-term policy decisions. The better the study the longer it takes, and consequently the less useful it may be. Policy-makers view evaluators as advisors, and feel no obligation to accept their recommendations. Furthermore, disagreement regarding evaluative criteria, i.e., explicit versus implicit goals of a program, sometimes contribute to nonuse of findings.

Cook and Lange (1975) suggest that educational and social policy be evaluated in four areas: the political process involved in the policy formation, policy implementation through specific programs, policy outcomes in form of specific products or consequences of programs, and the utilization of evaluation as feedback.

Utilization of evaluation results  Evaluation is only one of many inputs in the decision making process, political and bureaucratic considerations being others. According to Weinberger (1969), evaluation provides maximum utility when it can supply the structural reasons for the failure or success of a program.
Problems an administrator or decision-maker has in utilizing evaluation findings are generated by seven factors (Guba, 1975). Many different purposes may have been in the minds of administrators, evaluators, information sources, and other relevant groups. This will greatly confuse matters when it is time for utilization. The evaluation results may fail to meet the criteria of a good evaluation, which makes it difficult to make a meaningful application of the result. Discrepancies often occur between plans and operations of a program. Rapid fluctuation in the society itself requires rapid shifts in program goals and procedures leaving both policy-makers and evaluators outdated. Evaluators may be willing to produce whatever findings that are politically, economically, or psychologically advantageous, or become biased through close association with the project. A final source of difficulty stems from human and political factors. Human because people are involved with human characteristics and in power relations to each other, political because information often is taken as equivalent to power.

When administrators strongly believe in the effectiveness of a program or project, it is highly likely that an evaluation will be disregarded (Rossi, 1972). Interest is in positive findings, not in negative or no findings at all. Claims of programs made in public are ordinarily set at a higher level than anyone can reasonably expect to be attained, hence evaluation is threatening to the administrators. The more vaguely phrased the objectives of a program, the more difficult it is to obtain commitment to accept the finding of any evaluation effort.
Weiss (1972a) states that nonutilization of evaluation findings originates from the organization system itself and from the current state of evaluation practices. Organizations resist change because they are comfortable in status quo and negative findings may be seen as a threat to the organization. Evaluation results may not match the information needs of the decision-makers, or the timing might be wrong. Lack of clear direction for future programming may be another reason for nonutilization of evaluation results. The results may be ambiguous, implications unclear and leave a gap in translation of the results into programming. Evaluators are often reluctant to draw conclusions from their data. On the other hand, Weiss (1972a) comments that evaluation can add weight to the organization's own thrust for change, for better ways of operations. To increase the utility, evaluation may also pay greater attention to the organization's maintenance imperative that influences decision-making, and address the covert goals as well as the formal goals of the organization.

Weiss (1972a) points out that use of evaluation appears to be easiest when implementation implies only moderate alterations in procedure, staff deployment, or cost, or where few interests are threatened. Results that are a menace to the function of a total organization or occupational group are unlikely to be used by practitioners, or sponsoring agency, but must be used by policy makers who allocate funds. Even if results of evaluation are not used immediately, over time they can have cumulative impact on the climate of opinion and the legitimacy of programs.

Evaluators often disregard organizational issues, which are vital to the practitioners (Caro, 1975). Practitioners are interested in keeping
their jobs and going ahead, hence are apt to invite evaluation of only the most promising of programs. Evaluators are sometimes seen as inspectors, and therefore practitioners are reluctant to give information, or exaggerate reports of efforts. "Creaming", or a tendency to bias client selection in favor of those with a favorable prognosis, is often used by practitioners when they are judged by the management on client outcome. Evaluation is costly, and organizations may be reluctant to use money this way when it is needed to develop the program itself.

*Tension between administrators and evaluators* Tension often occurs between evaluators and administrators because of different basic orientation (Caro, 1971, 1975). In innovative projects, the stress may become even stronger because of the heightened risks involved (Brack, 1975). Caro (1971, 1975) points out that administrators are concerned with service and application of knowledge, the researcher with research and new knowledge. Administrators' interest is focused on solution of immediate, specific problems, whereas researchers often concentrate on long-range problem solving that can be generalized in time and space. Administrators may not want to explicitly state objectives and strategies, thereby avoiding firm commitment. Evaluative research needs clearly stated objectives and strategies in order to be meaningful. The researcher may be committed to scientific decision making, the administrator may have confidence in intuition. Evaluation attempts to discover inefficiency and to encourage change. Administrators and organizations tend to defend their programs and to resist change. Evaluators and administrators frequently emphasize different explanations for the failure of programs to
achieve its goals. Administrators tend to accept the validity of the theoretical premises on which their programs are based and to attribute failure to inadequate resources. Evaluators, being free to question program premises, are more likely to attribute failure to inadequate understanding of basic problems.

According to Caro (1971, 1975), evaluators approach social action programs from an academic perspective and are likely to underestimate the impact of political constraints, budgetary problems, and limitations of personnel and facilities. The practitioner may have difficulties in understanding the evaluator's emphasis on methodology. The researcher is searching for innovative alternatives to what the practitioner may regard as sound practice. The innovation of interest to the practitioner is likely to be viewed by the researcher as insignificantly incremental.

In order to minimize the tension between administrators and evaluators, Voht (1975) advocates a prior explanation of program objectives to all levels, as well as operational measures of these objectives. The latter need to be acceptable to both program administrators and researchers. Decision rules to be used in analyzing these operational measures and anticipated consequences would also have to be explained to all concerned before the evaluation research is undertaken.

Selection of criteria for program success A constant problem for an evaluator is to select criteria for determining the success or failure of a program.

Program objectives and goals are often used as criteria. Evaluation necessitates clearly specified, measurable goals. To specify objectives
involves assigning weight to different objectives (Cain & Hollister, 1972). Difficulties encountered stem from the need to reach consensus on the objectives, and from the complexity of the program outcomes. Consequences will be of both short-term and long-term nature as well as going beyond the more easily specified economic ones into those involving human behavior and attitudes.

One of the problems in program evaluation is that goals often are stated very vaguely or grandiously, i.e., to improve the quality of life, for which there is no broad consensus or indication of success (Bernard, 1975; Rossi, 1972). As a consequence, it is difficult to define program tasks in terms of changing individuals and organizations. One reason why program goals are vague is because they are matters of policies. To serve this political function they must be ambiguous, very general and acceptable to everyone (Orlans, 1973; Voht, 1975). To set a narrow goal is to risk attaining it, and thereby terminating the program. Narrowly defined goals may also misrepresent the actual aims of broad-aimed programs (Weiss & Rein, 1970). Voht (1975) gives additional reasons for ambiguity of goals in community development. One is the distinction between process and content objectives. Content objectives are concrete, frequently physical, and the result of community development activities. Process objectives involve the ability of the community to solve its problems and make decisions collectively, which conventional evaluation techniques cannot measure. Another reason is that community development seeks to achieve institutional change, rather than aggregate changes in the
behavior of individuals, which can be measured only by qualitative methods rather than by comparative techniques.

Evaluators who complain of absence of clear and concise statements of aims have not grasped the diverse and conflicting nature of social action programs, and therefore produce unrealistic, constrained views of program aims (Cohen, 1970).

Evaluation must focus on those criteria on which the customer wishes to base his choice (Coleman, 1969). Unless appropriate criteria are used, the result is irrelevant to the choice that must be made. The customer is often not fully aware of the most important criteria, hence the evaluator must discover these, rather than focus on those he believes to be most useful. Whatever criterion of program effect one might imagine, it could not intelligibly be evaluated in the absence of data which describe the character of the program (Cohen, 1970).

Generally, the objectives of the program are policy-matters and as such, are not within the province or the researcher to evaluate (Weinberger, 1969). The policy-makers decide upon the value of the program's objectives, on how much is good. Stake (1967) questions whether the evaluators should make judgments, although he believes they can process the judgments of others. Alkin and Fitz-Gibbon (1975) hold that the evaluator, even though he does not provide the judgment, must help the decision-maker to get a sense of meaning of evaluation outcomes. The decision-maker, who must put a value on outcome measures, has to find a way of interpreting them, and to get a sense of magnitude of effect. The latter must be understood before valid value can be attached and predic-
tions made. Education has relied on establishing direction of effects, but ignored the problem of establishing the magnitude of effects.

Whitaker (1974) takes the stand that it is necessary to allow the people being served by a program to provide standards for evaluation. Effectiveness is only one aspect of program quality, others are responsiveness and equity. Responsiveness, as used by Whitaker (1974), measures whether the public is better off because of the program, and equity the degree to which the program effects are allocated among various publics.

**Weak effects of programs** Many evaluation studies show weak or no effect of program efforts.

Bernard (1975) claims that one of the reasons why social service delivery programs usually fail is because these programs are directed toward changing role performance through changing behavior, attitudes, or values, and are often used as an effort to rapidly solve a major social problem in a crisis atmosphere. If these efforts to change people are expected to raise living standards for large numbers of people, or if such increases in income are expected to change attitudes, no program will be successful.

Campbell (1970) states that reformers and administrators achieve their permission to innovate by over-promising the certain efficacy of their new program. Programs representing large inputs of resources are likely to produce unanticipated consequences whose importance may rival, if not outweigh the intended aims (Weiss & Rein, 1970).
Assessment studies show that most action programs do not change the behavior of individuals or groups (Weiss, 1972a). Rossi (1972) finds that social action programs, geared towards disadvantaged audiences, usually show a weak effect. It is more difficult to demonstrate effects in programs directed towards changes in people rather than to technological solutions to problems. It is more complicated to reach and train the disadvantaged person in comparison with the less disadvantaged to achieve the same effect. It requires more effort to do so, and the cost is therefore greater. The more social services already supplied, the more difficult it is to add to benefits by supplying additional services.

Voht (1975) gives the following possible explanations for the weak effects of community development programs. The level of financial support is low in comparison to expected effect. These programs are expected to have a multiplier effect, that is, one change to instigate other changes. The efforts of a community development program are small, compared to other influences upon program participants and institutions. Program goals are often subtle, requiring changes in behavior for large numbers of people. Such goals can be only crudely measured, and, furthermore, are costly to assess.

Bernard (1975) mentions as one source of program failure the expansion of a small, widely known "successful" program. The expanded program fails because of lack of trained personnel, because new clients are different from the original group, or because high-quality services are expensive and expanded programs mean a reduction of staff/client ratio and
hence lower the quality. Finally, early claims of continued success are soon disputed.

As a result of the weak effect, a program may be improved, but may also be prematurely abandoned (Weiss, 1972a). It can also lead to abandonment of evaluation. To prevent this, Weiss (1972a) suggests less stress be placed on an over-all impact evaluation and more emphasis on a comparison of the effectiveness of varying conditions within the program, thus starting to identify elements and sub-elements associated with more or less success. Rossi (1972) proposes that in programs with weak effect, more high powered evaluation methods are needed to demonstrate effectiveness.

Methodological problems Most authorities do not use the term evaluation and research synonymously. Nor is the methodology of evaluation identical to the methodology of research (PDK Committee, 1971), because the purpose of research is to provide new knowledge that is universally valid; whereas, the purpose of evaluation is to provide information for making educational decisions, or information that is valid and useful within the decision-making context.

Evaluation research is viewed by Weiss (1972a) as a specific method of evaluation, using research methods in order to make the judgment process more accurate and objective. What distinguishes evaluation research and basic research is not method or subject matter, but intent, or purpose for which it is done.

Suchman (1969) considers evaluation research as the use of scientific method for collecting data concerning the degree to which some specified
activity achieves some desired effect. The significant difference between basic or nonevaluative research and applied or evaluative research is one of purpose and not of method. Both types of studies attempt to utilize research designs for data collection and analysis based upon the logic of the scientific method. The evaluative study applies this model to problems which have administrative consequences, while nonevaluative research is more likely to be concerned with theoretical significance.

Voht (1975) maintains that evaluation research is fundamentally the same as basic social science research, the major difference being the variables involved. Evaluation research is concerned with policy relevant variables, some of which are manipulable.

Rossi (1969) concludes that whatever differences there are between pure research and evaluation research resides primarily, if not exclusively, in the social and political relations of the research processes involved. The differences lie in the kinds of organizational context in which typically the one or the other type of research is carried out and in the relationship among researchers, those who provide the funds for research, and the audiences to which research findings are directed.

Alkin and Fitz-Gibbon (1975) state that the setting of the evaluation study firmly grounded in the decision framework is the crucial difference between evaluation and research.

Research methodology is usually interpreted to be the one used in psychology; the experimental model with its design of pre-test, treatment, post-test, and comparison with a control group. This model is the only one that produces "hard data" which permit the drawing of cause-effect
conclusions. The classical research method has been proclaimed by many as unsuitable for educational and social action programs (PDK Committee, 1971; Borus, 1972).

The PDK Committee (1971) gives several reasons for rejecting the experimental design. The laboratory used in classical research provides a context-defined environment in order to produce universally valid, new knowledge. Programs, on the other hand, are conducted in field-settings that approximate or equal reality, in order to make possible judgments about some phenomena. A second reason is the effects of interventions. The experimental researcher has some degree of control over variables and manipulates these in some way to study their interaction. The evaluator has little or no control over variables and wants information about actual situations. A third reason for rejecting the experimental method is the terminal availability of research data, i.e., information is produced after treatment is completed. Evaluation is concerned with optimizing improvements in education which requires continuous feedback. Research tries to ascribe an effect to a cause. Evaluation is usually involved with multiple assessments simultaneously. Research strives for homogeneity in all variables except the one tested and is able to control sources of error variance. Evaluation research lacks this opportunity to control variables by random selection and assignments.

A sixth reason is the inapplicability to evaluation of assumptions underlying statistical analysis appropriate to research. Normality of distribution, random sampling, and analysis of variance with its additivity assumptions, are not likely to hold in typical evaluation situa-
tions. Research usually has restricted decision rules and is limited to rejecting or not rejecting a null hypothesis, and to judge one treatment better than another. Evaluation has broader decision settings.

Finally, basic research keeps treatment constant whereas evaluation is designed to promote dynamic development of programs, to refine them continuously.

Notwithstanding the above reasons for rejecting the experimental methodology, the experimental design has its place in program evaluation. Weiss and Rein (1970) feel that it should be used more often than is currently done. If one aim of a program or a single objection to a program is important enough to justify collecting data which will lead to a relatively unquestionable conclusion, and when the program can be given the form of standardized replication within relatively controlled situations, the experiment is legitimate. Suchman (1970) advocates the experimental design for evaluation of model programs. Alkin and Fitz-Gibbon (1975) observe that the experimental design will detect program results that a purely descriptive evaluation could not validly interpret. Cohen (1970) claims that the chance for success of experimental approaches to social action will be directly related to a program's political independence, its specificity of aim, and its fiscal strength. Because of political constraints on social action program evaluation, his advice is to experiment only when the substantive issues of policy are considerable and reasonably well defined. Weiss (1972a) claims that the experimental design is best suited when the purpose of evaluation is outside the immediate program. Decisions on continuation or abandonment of the program; on whether to
advocate nation-wide use of the program model require great confidence in
the validity of the research and therefore experimental design.

Sjogren (1974) contends that the variety of data and information
needs in educational institutions requires the skills, methods, and in­sights of not only psychology but the social sciences and the humanities
as well. Education is an interdisciplinary endeavor and evaluation
efforts should also be interdisciplinary. No single style or method is
all-sufficient for educational evaluation. Each style or method provides
unique information and insight. A comprehensive evaluation activity would
gather data in different styles and report information of importance to
each of several audiences.

Evaluation research depends heavily on external conditions and must
make some compromises of design and methodology (Jones & Borgatta, 1972).
The situation in social action and broader educational programs is
virtually uncontrolled. A multitude of variables other than the program
under evaluation might be operative and effecting the particular goal
being investigated, i.e., changes in the national economic or political
situation, other agencies providing the same program.

Control-groups The experimental design makes use of control
group(s). This group does not necessarily receive any program. Usually
the choice is between the experimental treatment and the standard treat­
ment. If no treatment, a "placebo" program reduces the possibility of the
Hawthorn effect. Control groups may be contaminated through association
with people in the program or by receiving the same treatment from another
agency. To compensate for these and other variations in factors not in
the treatment, random assignment to treatment and control group is used in the experimental design.

Program evaluators, on the whole, tend to agree that some form of control group is essential when the effect of a program upon its participants is assessed. Is there any difference if a program exists and when it does not? is the question to which they seek an answer (Borus & Buntz, 1972; Cain & Hollister, 1972; Cohen, 1970; Hardin, 1972; Voht, 1975; Weiss, 1972a,c).

The choice of control group can affect the impact estimates of program effect, which statistical control can counteract, but not counterbalance. Yet randomization is seldom used in social action programs. The major obstacle to randomization lies in the political and administrative resistance to the idea of experimentation and to the controls over programs and projects which experimental design makes necessary (Rossi, 1972). Practitioners want to assign people on basis of their professional knowledge and experience, and oppose randomization because they are not free to administer "optimal" treatment to the control. Another ethical problem is that an assignment of control infers inferior treatment. Nobody wants to be a control (Caro, 1975).

There are alternatives to randomized control groups. Before-and-after studies assume that each subject is his own control and that the before-measure is the behavior if no program occurred. This method is likely to create bias in impact estimates. The before-measure might be taken at a point when the participants are at a lower state than is normal for the group, i.e., reflect a transitory condition (Cain & Hollister,
Structural changes in personal situations, i.e., permanent displacement from a job, before and after the program might influence the measurements at both or either points without having anything to do with the program. Significant environmental changes such as seasonal and cyclical fluctuations in labor markets might also enter in. The principle of regression towards the mean dictates that even in the absence of a program, groups would have greater-than-average improvements in their work situation.

Drop-outs from programs are often used as control groups. As this group is affected by their incomplete treatment, it is not good as control for those completing the treatment (Hardin, 1972).

Matched control groups, or comparison groups, is an attempt to make the control group as similar to the treatment group as possible, matching variables such as age, economic status, education. One difficulty with this method is that it is almost impossible to define the characteristics on which people should be matched; to know which ones will affect whether the person benefits from the program or not. Another hardship stems from difficulty in locating the comparisons, especially in poverty research areas.

Self-selection is another problem. People who choose to enter a program are likely to be different from those who do not. Hardin (1972) and Borus and Buntz (1972) advocate random assignment of qualified, interested persons to treatment and control groups as qualified, non-enrollees are more appropriate a control group than the general target population.
Alkin and Fitz-Gibbon (1975) suggest that when all participants are included in a program, the people take turns in getting treatment by switching the treatment midcycle. Participants would be randomly assigned to groups, which would act as treatment and control alternately. Contamination is a serious problem in this design. Alkin and Fitz-Gibbon (1975), using the argument that "most needy" is an unreliable measure in itself, recommend that those people a practitioner deems "most needy" be assigned to treatment, but that borderline cases be randomly allotted to treatment and control.

Borus and Buntz (1972) propose that if random selection of participants is not feasible, multivariable techniques be used to account for as many possible differences between participants and control group members as are measurable.

Random assignment or matched controls usually imply a small group study. There is a question about how representative a small group is in terms of participants and characteristics of the program. Cain and Hollister (1972) tentatively suggest that the small group study is appropriate for those evaluations that have a design concept which is radically different from existing design or where there is a quite narrow hypothesis which requires detailed examination. Where the concept underlying a program is quite broad, and where large amounts of resources are to be allocated, the large group approach might be the better choice. Here data are gathered over a wide range of environments and control over the characteristics of participants and nonparticipants is achieved by statistical methods rather than randomization or careful matching of individuals.
Wortman (1975) is concerned that lack of adequate control allows statistical regression to become a plausible alternative explanation for observed differences between treated and nontreated groups. For groups artificially matched on extreme scores, a typical situation since the neediest usually are given such social action programs, i.e., Headstart, the regression will not only be toward separate means, but the amount of regression can be substantial and thereby produce significant differences for groups with widely differing means and/or small standard deviations. Extraordinary care is hence necessary in evaluating the success or failure of policy-relevant projects, especially those of national importance.

Validity problems in program evaluation A multitude of problems limit the validity and usefulness of manpower program evaluations (Borus, 1972). Issues arise from determining the program unit to be studied, in the choice of proper control group, when determining the optimal method of securing information on these effects of the program, in the choice of variables to measure program effectiveness, in the calculation of cost, in the long-run projection of benefit estimates for policy use, and in the measurement of secondary effects of manpower programs.

Wortman (1975) cautions about problems with quasi-experimental design in regards to internal, external, construct, and conclusion validity. The internal validity refers to the inferential process involved in the assignment of causality for the observed effects to the treatment or social action program. External validity is concerned with the generalizability of that causal inference to other populations, settings, or treatment variables. Construct validity refers to confounding the cause-
effect relationship with another or improper translation of theory into either dependent or independent variables. Conclusion validity is a subset of threats to internal validity, which stems from errors introduced by small sample size, differences in treatment administration, unreliable measuring instruments and statistical tests involving this information. Worthman (1975) considers formative evaluation more of an art than science. It can reveal important differences between theory and treatment that lead to changes in construct validity and in theory. Summative evaluation assumes that the treatment has been properly implemented and is concerned with how effective the program is in attaining its objectives. A successfully operating program or a treatment that achieves worthless goals cannot be evaluated positively in summative analysis.

There is a clear temporal order contained in the evaluative components but the ordering among the various forms of validity differs for the theoretician and the applied researcher (Worthman, 1975). The theorist prefers the order of internal, construct, conclusion, and external validity; the practitioner the one of internal, external, conclusion and construct validity.

**External versus internal evaluation** One of the controversies in program evaluation centers about the use of external and/or internal evaluators. As both types of evaluators have strong and weak sides, a decision as to which kind to use will have to be made in each specific situation. Factors to take into consideration include the intended use of the evaluation, autonomy, objectivity, understanding of the program, and administrative confidence (Weiss, 1972a).
The external evaluator, with his prestige and authority, may be able to get the administration to utilize evaluation findings in a policy context. Administrators might have higher confidence in the competence of the external evaluator. The outsider, because of less involvement in the program and having less at stake in the outcome of the evaluation, is generally viewed as being more objective than are inside evaluators. The outsider might have less access to information about what really goes on in a program than the insider, and less awareness of issues facing the program and organization. On the other hand, he may be able to take a wider perspective on the program than the insider, who may be prone to limit the evaluation to the basic assumptions and framework of the program.

Wrightstone (1969) proposes that an external evaluator be employed for controversial and sensitive issues and decisions regarding them, as well as when an internal, trained evaluator cannot be found.

Steele (1975) suggests that after a time period when experts tried to take evaluation out of the hands of program people in an effort to increase objectivity, the trend now is to involve as many people as possible.

Evaluation of Extension and Adult Education Programs

Traditionally, program evaluation in extension has focused on two aspects, namely on the number of participants in the specific program, and on changes in practices on part of the participators. The assumption underlying the counting of the audience is that if a program does not meet
the need of the participants, they will not attend. Adoption of a new practice is an indication of a successful educational effort, i.e., a change of behavior, but emphasizes the end result, or product, of a program.

Following the general trend within education, extension personnel have broadened their concept of program evaluation. For example, Steele (1970), stating that evaluation needs to be purposeful and to contribute to present and/or future programs through being an input to the program planning process rather than an end to be achieved in the process, proposes the following definition of program evaluation:

Program evaluation is the process of judging (or a judgment as to) the worth or value of a program. This judgment is formed by comparing evidence as to what the program 'is' with criteria as to what the program 'should be' (p. 8).

According to Steele (1970), evaluation has three essential elements: criteria, evidence, and judgment. Often what has been called program evaluation is merely a program description, in which there is evidence of results, but no indication of either criteria or judgment of the program. With only a description it is difficult to really use the evidence in a way that is meaningful in program decisions. Thus, the areas where extension in general and many individual extension workers in particular need to expand their concept of evaluation are those of criteria and judgment.

The program planning decision that needs to be made identifies the focus of the evaluation. Programs can be judged in terms of several different characteristics. At least five are usually assessed: quality,
suitability, effectiveness, efficiency, and importance. How the evaluation will be used should determine the characteristics that will be examined. Steele (1970) proposes that when resources are limited, the most important characteristics for examination are those of efficiency and importance. A set of criteria has to be developed and selected, both at macro and micro levels, that will be of most use in getting the answers that will help in making decisions about programs. Groups, advising on program planning, need to be involved in the evaluation as well.

As seen by Steele (1970), there are several tasks ahead for evaluators in extension educational programs. Among these are to develop criteria related to the different characteristics of a program, and to understand what characteristics are used as bases of judgment when a particular kind of criteria is applied. Others are to identify which criteria are most important in which situation, and to acquire increased experience in making judgment of various characteristics and of the quality of those judgments. It is lastly suggested that not only is it necessary to find out how programs measure upon quality, suitability, effectiveness, efficiency, and importance, but also how the degree of "goodness" on each affects the other and the total value of the program.

Evaluation of programs aimed at low-income adults

Adult educational programs have always been a part of the American educational system, but have reached mainly the middle class of the population. During the last decades, efforts have been made to include other strata of the American society. Federal legislation, e.g., Part F of Title I of the 1968 Amendments to the Vocational Act, provides funds for
home economics programs with greater consideration to social and cultural needs. Conjointly with the funds comes the requirement for program evaluation.

The Expanded Food and Nutrition Education Program (EFNEP) is one answer by the Cooperative Extension Services to the call for including low-income clients in its programs. The operational objective is to improve families' nutritional intake and practices. It uses the person-to-person approach through paraprofessional aides, recruited from the neighborhood. Most EFNEP started during 1968 and 1969. As the use of paraprofessionals was new in extension, much attention has been given to the evaluation of the program. The greater part of these appraise the extent to which changes in the clients' dietary habits have occurred as a result of participation in the program.

Assuming that evaluation should be logical, empirical, and objective, Elliot (1972) suggested a modified Stake evaluation model for the assessment of the EFNEP. The evaluation begins with a statement of the program rationale, or philosophical background and basic purposes, from which the evaluation plan is formulated. Both stated and unstated goals of educators and learners are included in the appraisal and examined in relation to participants and paraprofessionals.

As antecedents, or conditions existing prior to learning, data on the learners, the paraprofessionals, professionals, and the programming situation in the county in question are to be collected. Elliot (1972, p. 7) proposes information be compiled about the economic characteristics of the learners and paraprofessionals; food consumption patterns of
learners and paraprofessionals; social and physical environment of the learners; the indigenous or ubiquitous paraprofessionals, and their level of nutrition knowledge; the experience of professionals in this type of program; and the staffing patterns in the county.

A wide variety of transactions, or activities in which learners and teachers engage together with the instructional material used, need to be examined, and Elliot (1972, p. 7) suggests the following be considered: relationships between paraprofessionals and families, paraprofessionals and professionals, and professionals, paraprofessionals, and volunteers. The learning situations provided for paraprofessionals and for professionals; how paraprofessionals use teaching materials and the ways in which they work with families-individuals-groups; and the involvement of volunteers in the program require attention, as do the channels for decision-making within the program and communication links between paraprofessionals, professionals, community influentials, other community agencies, co-workers, and families. Consideration about when it is appropriate to move families to another phase of the program needs to be given. Both intended and observed outcomes of the program are recorded and taken into account in this evaluation model.

Some data about antecedents, transactions, and outcomes are available in organizational reports and internal data systems. Elliot (1972) seems to advocate the use of structured interviews, questionnaires, and case studies for gathering additional data, as well as content analysis of logs written by aides after each visit to a family.
Content analysis, correctly executed, is a research technique for objective, systematic, and quantitative description of the manifest content of communication. Applied to the logs, it might be possible to indicate behavior changes in the clients, as well as to identify gaps in subject matter content, in attitude, and in methods used by the paraprofessional.

As evaluation design, Elliot (1972) proposes a variation of the longitudinal design. Instead of assessing the same subjects at different points over a time span, different populations at different stages of the program are involved at one point of time. The assumption is that these successive cohort groups are representative of the same target population.

Because the process of defining criteria for making judgments about the program can further clarify goals and intents, as well as promote an examination of the relevance of values to the situation, Elliot (1972) suggests that both professionals and paraprofessionals be involved in this.

Although Elliot (1972) envisages that judgments about the program will lead to program improvement and accountability, she is vague about who is to make this judgment once the evidence is collected and criteria identified.

The evaluation itself needs to be evaluated in order to make it possible to improve evaluative efforts. Elliot (1972) proposes the use of the scientific, practical, and prudential criteria discussed by the PDK Committee (1971) be applied.
Elliot (1972) is concerned with evaluation as opposed to evaluative research; with "what is" as opposed to "why". She is also interested in the examination of a specific program in a specific community where the question of generalizability is of low priority.

The most questionable in the proposed evaluation design is the assumption that successive cohort groups are representative of the same target population. Low-income clients are a heterogenous group; the recruitment technique used is out-reach by the paraprofessionals. The number of participants is usually low, which excludes statistical correction of uncontrolled variables.

The Synectics Corporation evaluated the program performance of the Expanded Food and Nutrition Education Program during April 1970 through March 1971. This was the second in-depth evaluation of this program, the first being conducted by Datagraphics during January through July, 1969. Objectives of the study were to validate the earlier study, as well as to characterize program activities and achievements as related to program variables; to assess new dimensions or emphases of program operation; to identify and assess problems and barriers encountered beyond the formative stages of the program; and to identify the need and/or opportunities to expand or modify the program beyond its current scope.

The same program units used in the 1969 study were utilized with three added units where special youth activities had been underway since the initial stage of the program. All study locations were selected by the Cooperative Extension Service, United States Department of Agriculture (USDA), to include different geographic areas, urbanizations, ethnic back-
grounds of personnel and participants, and to represent variation in program design and operation rather than to provide a representative sample of the national program.

Data were collected from information available through regular program sources as well as from personnel associated with the program, participating families, youth participants, and volunteers. Additional information was sought from state extension personnel, other local extension personnel, and from representatives of other relevant agencies within the community.

The analysis of data, as described in the report available for review, consisted of review and analysis of field notes and other nonsystematic data; of routine data processing for description of the study population; of analyses to test specific hypotheses related to overall program, and of special analyses to explore specific issues arising out of direct observation and special questions posed.

A descriptive model, representing components of the EFNEP and their complex interactions was prepared, principally to be used as a management tool. Implementation of the model provided the primary mechanism for the development of analyses supporting the investigation.

The major analysis, through the use of correlation, examined the relationship between the set of program descriptive characteristics and the criterion variables for identification of significant characteristics of program operation, personnel, and participants associated positively or negatively with program performance.

The analysis of the adequacy of the family's economic resources to provide an adequate diet consisted of comparing reported monthly food cost
of the family with the projected USDA Low-Cost Food Plan data for September 1970. The latter data were also used to project the total family economic resources required to remain out of poverty.

Conclusions of the study were generalized to the overall program, not necessarily applicable to all locations. They dealt with family nutrition practices; nutrition-related practices; the indigenous aide; the supervising agent; program leadership; the programs; ability to reach target audience; operational structures; youth activities; community role; and the need for remedial action.

The evaluators reported that nutrition education objectives were met, but fuller effectiveness could be achieved with a more precise allocation of the aide's and the program's effort. The program promoted improved nutrition related practices such as: money management; food purchasing; storage, handling and preservation of food; sanitation procedures; meal planning and use of utensils; and generalizable skills in food preparation. The indigenous aide was appropriate and effective in nutrition education of low-income families. Reasonable progress had been made in establishing the program within the total community setting, and the program reached the target population. Nevertheless, the link between homemaker and youth efforts needed to be strengthened, and although the supervising agents were dedicated in their program role, their capabilities were increasingly strained as the scope of the program expanded. On the whole, this second evaluation substantiated and extended the need for remedial actions, identified in the first report, but in a more urgent and
less optimistic tone, as the first evaluation capitalized on the program's strong sides.

The report noted a growing concern for program stability at both local and state level, as well as for meeting future program support needs. It also called for improved utilization of program resources.

The evaluators also commented on the records that are sent from local level to state and federal levels in a highly consolidated fashion, a form that makes sophisticated analysis of information impossible. The local levels most often lack facilities and capabilities to perform the analysis itself. Formal efforts to relate cost and benefit data at the local level were rare, which the evaluators attributed to a possible fear on the part of program leaders at all levels that the cost/benefit data would force decisions about the program. The evaluators consider the 24-hour food recall and the number of families reached the most promising source for cost/benefit data at the time of the study, possibly related to some intermediate performance measures such as aide/participant contact hours.

A study outside the EFNEP program but with most features in common, was done by Nelson and Jacoby (1973), who assessed programs aimed at low-income adults in New York. The study was the second in a series related to the evaluation of federally funded adult consumer-homemaking education programs.

The purpose of the study was to make a formative evaluation of the program, i.e., to provide information about the impact of the program on the participants, and paraprofessionals, and to provide information for decision-makers needed for deciding on program recycling or revision.
The research design used was a pre- and posttest one, using program drop-outs as a comparison group.

Ten program sites, nine urban and one rural, were selected by the Bureau of Home Economics Education, New York State Education Department, as representative of fulltime adult consumer-homemaking programs operating in New York.

The programs carried out at the different sites were tailored to local needs, following certain state guidelines. In order to provide the flexibility necessary in a program where the participants help to set priorities, the investigators developed a bank of objectives from which staff could select those performance objectives appropriate for their program. Content validity of the bank of objectives was established in the study.

A sample of 108 participants was drawn through a process of random, unannounced visits by the investigators to the centers. The number of participants from each center was weighted according to the previous month's attendance.

All of the 26 aides who were employed in the ten centers at the time of the study participated. Ten newly hired aides employed in similar programs made up a comparison group.

Data were collected using such nontargeting techniques as interviews, systematic observation, unobtrusive ratings by teachers of participants and paraprofessionals, and recording of critical incidents by local staff and investigators. Field staff kept comprehensive records of attendance, resources, and program content.
Questions regarding demographic background were not asked directly but gained from monthly teacher reports, observations by observers, and the Background Information form completed by teachers at the conclusion of the study. Demographic data included race, age, educational background, marital status, health, income, source of income, number and age of children. Profiles of the typical participant and aide were described and compared.

Several instruments were developed and used in the study. To measure participant progress toward cognitive, affective, and psychomotor objectives of the program, the following instruments were used: the Participant Interview Schedule, a form for recording critical incidents, and the Consumer-Homemaking Descriptive Rating Scale.

Instruments measuring attitude and proficiency of professional staff were a structured form used by the investigators during systematic observation, the Teacher Interview, Resource Record, and a form for recording daily programs.

Instruments for gathering information about the attitude and proficiency of paraprofessional staff were a scale assessing attitudes toward paraprofessional jobs in human services, a descriptive rating scale for judging work performance, and an aide interview/questionnaire. Additionally, the Participant Interview Schedule was used to measure the impact of the instructional program on the paraprofessional staff.

Information about processes and facilities was gathered from an attendance register, the Background Information sheet, the Child Care Record, and the Inventory of Facilities and Equipment.
Content validity and reliability were established during the study for the Consumer-Homemaking Descriptive Rating Scale, the Rating Scale for Evaluation of Paraprofessionals, and the Participant Interview Schedule.

Program effectiveness was measured in terms of progress of program participants toward cognitive, affective, and psychomotor objectives, acceptance of program by participants, progress of paraprofessional staff, progress of professional staff, and processes and facilities. The t-test determined the significant differences between gainscores of participants attending at least 15 times during 6 months, and those attending less often. Where indicated by small number of subjects, the nonparametric signed rank statistic was used, when comparing means of variables for the ten centers, or looking at individual programs. A correlation matrix was used to investigate relationships between the participant interview and descriptive rating scale, demographic data, and hours spent by participants in various center activities.

An Index of Program Effectiveness was compiled from rankings of the ten individual programs on Consumer-Homemaking Descriptive Rating Scale gain scores, attendance gain, use of resources, maximization of teachable moments, quality of daily preparation, and critical incidents. The Kendall's Concordance Coefficient was used to measure the degree of association among sets of rankings in the Index of Program Effectiveness.

This Index of Program Effectiveness, according to Nelson and Jacoby (1973), defines possible criteria for determining program success, with a statistically significant degree of association among rankings on the criteria (p<.02).
Characteristics of programs ranking high on the Index were examined to describe the successful consumer-homemaking program for low-income adults. According to this, such a program is characterized by city supervisor of home economics giving top priority to development and supervision of adult centers, direct teaching by professional home economists, varied program offerings, close cooperation with other agencies, planned educational activities for children, at least a rudimentary training program for paraprofessionals, and continuous recruitment of participants assisted by expanding outreach activities and supportive news media.

Santopolo and Kell (1973) tried to establish exactly what tasks the nondegree person can, does, or should assume in Cooperative Extension Service. The specific purposes of the study were to identify the on-the-job behaviors and attitudes of the paraprofessional aides; to determine critical job requirements by tabulating and classifying the identified behaviors and attitudes; and to draw implications from these data for the training of EFNEP aides.

Method used was an analysis of critical incidents acquired from 40 paraprofessional aides who had worked more than 3 months in the EFNEP program in selected areas of Kentucky. The aides, in taped interviews, were asked to recall a time when they felt exceptionally good or exceptionally bad about their work. No attempt was made to limit the respondent in terms of how long ago their incidents happened, whether they described the effective incident before the ineffective one, or whether their incidents were actually effective or ineffective as perceived by the investigator. The interviewer asked several questions designed to
elicit comparable data in the event the critical incident did not include these.

Before analysis, each incident was transcribed, reviewed and edited in order to guarantee anonymity and assure readability. Incidents were considered effective if the aide felt the client was making obvious progress, had maintained a continued interest in the program, and/or was actively participating in the program. Ineffective incidents were those in which the aide felt there was no evidence of improvement, the client was disinterested, or lacked commitment to the program. Of 80 recorded incidents, 66 were considered usable by the investigator. Of these, 36 were classified as effective, the remaining 30 as ineffective incidents.

Individual personal contact with the client was considered by the aides to be critical to the effectiveness of the EFNEP delivery service, as was the making of a series of follow-up visits, and the use of materials directed to the needs of the specific audience. Other factors related to effectiveness, as perceived by the aides, were the taking of initiative or decided action on the part of the aide with regards to real or imagined client problems of the moment; the gaining of the client's confidence; and the establishment of credibility. Enthusiasm, a generally positive attitude, and persuasiveness on the part of the aide were deemed vital to their success with clients, as were self-confidence, commitment to the client, and concern about the client's welfare.

Ineffectiveness was associated with lack of understanding and/or knowledge of specific client types and consequently the inability to relate or work effectively with them. Negative or discouraged attitude in
general, lack of self control, and a focus on perceptions about the client or client environmental conditions were other variables linked with ineffectiveness. The majority of the ineffective incidents stated that the end result was negative because the client did not seem interested in improving her situation despite repeated efforts on the part of the aide.

Based on their findings, Santopolo and Kell (1973) recommended that the training of aides should be conducted on a one-to-one basis, or in very small groups, as this is the manner in which aides interact on the job. The major training approach should consist of the integration of problem-solving experiences concerned with how to teach nutritional education and role simulation activities dealing with human relations skill development. As no program can be operationally effective without support of organizational personnel, any training program for nonprofessionals should focus attention on not only involving the trainees, but all levels of staff responsible for the success of the program. Suggested ways for the latter is to have organizational staff issue a written statement stating their approval of the training proposed, so that all other levels of staff are aware of their commitment and approval; to have appropriate personnel involved in helping to develop the training philosophy and/or the curricula; and to expose others in the organization to the content of the training of the aides, thus furthering better understanding of the aides and their work.

With the ultimate aim to propose guidelines for the selection, training, and supervision of paraprofessional personnel, Yerka (1974) studied the effectiveness of aides in the Expanded Food and Nutrition
Education Program in working with low-income families. In her view, the aides function as a bridge between the program and its clientele and hence are an important link in the system.

Twenty aides, employed in one New York City community, together with a total of 89 homemakers participated in the study. All data were gathered between the fall of 1970 and fall of 1972.

As the study was a third component of a larger project, several instruments and measures for appraising the effectiveness of paraprofessionals had already been developed and field tested.

Characteristics deemed influential to the effectiveness of the aides were the paraprofessionals' knowledge and understanding of the teaching-learning process; their attitude towards assuming their role; and dimensions related to job persistence and job performance. The criterion used to determine the effectiveness of the aides was the success of families in reaching program objectives. Four measures were taken, two of which dealt with nutritional practices, i.e., gross food consumption and the thereby attained level of nutrition. A third measure was of the homemakers' knowledge of nutrition, and the last their recognition of assistance desired or received from the program.

The research design utilized a nonequivalent control group which consisted of an experimental group of homemakers drawn from the program clientele, and a comparison group. As it was not possible to assign subjects randomly to the groups, the comparison group comprised of "naturally assembled collectives as similar as availability permitted" (Yerka, 1974,
p. 80). The difference between the experimental and comparison group was the length of time the respective homemakers had been enrolled in the program.

The 76 experimental homemakers had been taught by one paraprofessional participating in the study over at least an 8-month period of time. The comparison group consisted of 13 homemakers who had received only initial contact, defined as three or less visits from an aide. Measurements of the homemakers were taken before and after intervention, or limited intervention, respectively.

All employed and trained aides who worked with program clients over at least an 8-month period during the study were included. All aides were measured on three variables, and a subsample of 14 on a fourth measure.

Demographic data for aides were collected from their job application forms, and for the homemakers by the aides in their normal recruitment procedures.

Similarities and differences between homemakers and paraprofessionals were investigated through analysis of covariance and F-test for significance of adjusted group means.

A t-test for independent sample proportions for each of the variables from the demographic data was conducted in order to establish significant differences between the experimental and comparison groups of homemakers.

The Extension Service, United States Department of Agriculture, records were used as the source for information on the food consumption practices and nutrition knowledge of the families. The analysis of covariance and F-test examined the difference between the experimental and
comparison homemakers' scores on food consumption, nutritional level and nutritional knowledge. Interrelationships among homemaker practice and knowledge scores were calculated.

As the experimental homemakers did significantly and/or practically better on the food consumption, nutritional level, and knowledge scores than did the comparison ones, Yerka (1974) concluded that the paraprofessionals made an impact on the clients' lives.

An interview was conducted to judge the homemakers' recognition of assistance (attitudes/behavior) received from the program. External interviewers, deemed accepted within the community, were engaged, and the homemakers were paid for the time spent in the interview. The interview schedule used was field tested in both rural and urban settings.

The analysis of variance (F-test for significance and t-test for proportions) was used in analyzing the differences between the experimental and comparison groups of homemakers in recognizing program assistance received or desired.

The three instruments with reliability and content validity estimated in an earlier study by Stuhlmiller (cited in Yerka, 1974) included a 39-item interview-achievement test on teaching-learning principles and strategies. Responses were practice-oriented and used a three-point scale. The reliability of this instrument was .91. A 42-item Likert type four-point attitude scale with a reliability of .82, collected information about the aide's attitude toward her job and its environment. A third instrument rated the aid's job persistence and performance factors related to persisting behavior. It consisted of 15 items and used a three-point scale. Estimated reliability was .88.
Developed for the study was a 23-item on-the-job performance rating scale, as well as a form for analyzing the written family records (logs) kept by the aides. In the first instrument, the supervisor rated the aide using a five-point descriptive rating scale. Content validity was established through literature review and interviews. In a field test, inter-rater reliability for two pairs of judges was .99.

A descriptive profile of the paraprofessionals was developed from the demographic data, the aides' knowledge of teaching and learning principles and strategies, their attitudes toward their job, and their job persistence and performance behavior.

Intercorrelations among the paraprofessional instruments as well as between the aide's scores on these instruments and the criteria of success for program participants were determined.

The difference between the mean scores achieved by the paraprofessionals on pre- and post-evaluations of their knowledge about and attitudes toward their work were tested for significance by a t-test for paired observations.

Findings indicated that mature indigenous paraprofessionals with educational levels from 9-12 years had acquired the characteristics which made them effective in bringing about changes in the lives of their clientele. Their knowledge of teaching and learning was important to the success of families in the study, but their on-the-job performance ratings were most highly associated with criterion practice variables.

The attitude scores of the paraprofessionals, although generally changing positively during their employment, did not result in positive
relationships with criterion variables, except for the homemakers nutritional knowledge. On the other hand, as Yerka (1974) pointed out, nutritional knowledge does not ensure that food will be consumed, nor does lack of knowledge mean that certain foods will not be included in the diet.

A major determinant of whether or not the paraprofessional could influence the homemaker to change patterns of behavior was her ability to demonstrate that her teaching was of value to the homemaker. Once the client believed the aide understood and related to her, the aide became a "significant other" and was looked to for assistance and information.

Those aides who were observed to practice principles of learning and teaching were effective in motivating homemakers to practice their learning. The more knowledge an aide had about working with people and the better she felt about her job, the more able she was to impart knowledge.

Finally, stepwise multiple regression examined the effectiveness of the paraprofessionals, indicating which and how many of the independent variables were predictors of the paraprofessionals' success.

The study indicates that knowledge of how to teach and how people learn, plus skillful practice of teaching components were predictive of the success a paraprofessional may have with homemakers.

The Expanded Food and Nutrition Education Programs are often assessed in terms of changes in nutritional practices on the part of the participants. Other aspects of changes in the participants' food habits have been explored such as the durability of the change (e.g., Rountree, 1973) and factors related to the adoption of food practices (e.g., Roy, 1973). Verma and Jones (1973) investigated the relationship between the dietary
levels of homemakers participating in the EFNEP in Louisiana and the
length of their participation; the kind of learning experience, i.e.,
individual visits or group meetings, or a combination of these; and the
intensity of the learning exposure, i.e., number of visits and/or
meetings.

The study was conducted from February 1970 to January, 1971. Of the
35 parishes (counties) in Louisiana that had an EFNEP program, 15 took
part in the study. Each aide in these parishes prepared a list of all
homemakers with whom she was working. From these lists the homemakers
were randomly assigned to one of three treatments. Data were collected
from the family and food record forms used in the EFNEP program. After
benchmark data were collected, six additional food records were collected
at 2-months intervals by the paraprofessionals under the supervision of
their regular supervisor.

Data were examined through one-way and nested analysis of variance to
test for significant differences in the number of servings of the four
food groups consumed by homemakers over the several time intervals and
types of learning experiences. Linear correlation investigated the asso­
ciations between the number of servings of the four food groups and the
intensity of the exposure to the different learning experiences.

Consumption levels of all four food groups increased during the edu­
cational program, especially in the milk and fruit/vegetable groups. The
increase occurred early in the program, reached a plateau and, generally
held that level. The authors, therefore, suggested that emphasis be
placed on continually contacting new homemakers rather than working
indefinitely with the limited few initially enrolled in the program. But as other studies have shown that those who are dropped from the program tend to revert to old food habits within 6 months, the clients should not be completely dropped but the intensity and character of the contacts be altered after a reasonable time period. Changes proposed include more group meetings, fewer and shorter home visits, greater use of volunteer leaders, and more use of mass media and mailed teaching aids.

Visits and group meetings had about the same effect on dietary change. Nevertheless, Verma and Jones (1973) recommend that both methods should continue to be used, as more people can be contacted through meetings, but home visits are important in creating and maintaining client relationships and in helping to solve problems.

In program evaluation it is important to look for unintended as well as intended effects of a program. The investigator has only found one study related to side-effects of the EFNEP program. It was conducted by Ramsey and Cloyd (1975). Most of the aspects mentioned in this study are in keeping with the objectives of the EFNEP program, i.e., food-related behavior defined as general attitude toward food, meal planning, shopping pattern, and dietary changes. The second type of side-effects was community-related behavior, described as sociological and sociopsychological involvement in the community and measured in terms of feelings of powerlessness, changes in problem-solving method, and social participation. The differences between program participants and a group-matched comparison group were slight, if any. Given possible explanations for the absence of greater evidence of success were the short period of time
between measures (one year), difficulties in changing cultural patterns, and changes in the larger society.

Evaluation of Family Planning Programs

Family planning in terms of contraception has been practiced by individuals throughout history, but organized efforts to provide the general public with knowledge, motivation, and effective means to plan their families is of more recent date. In the Western world, efforts frequently started by the endeavor of a few individuals, spread through private organizations, and from there gained national attention. A growing world-wide concern with overpopulation, especially in relation to the limited resources on earth (Meadows, et al., 1972; Mesarovic & Pestel, 1974; Presidential Commission on Population Growth and the American Future, 1972), helped pave the way for programs on a national basis. These were begun in India and Japan in 1952 (Rogers, 1973) but it was not until the 1960's that national family planning programs came to prominence, and in many parts of the world, organized family planning is introduced through these programs. A research and evaluation unit is often established to guide the national family planning program, and to measure its results. Hence most program evaluation efforts are done at this level.

After reviewing evaluation studies of family planning programs, Reynolds (1972) found that program performance has received a good deal of attention, but that the methodologies still were controversial and the results inconclusive. The two major shortcomings of family planning
evaluation were the general failure to account for intervening variables other than those offered by the program, and the failure to account for efforts of the private sector. He saw a need for the development of clear concepts and for defining key terms; for improving existing data collection and analysis procedures; for refining survey instruments; for refining and validating measurements; for designing, testing, and installing evaluative systems, and for conducting evaluative studies that will provide information that can be used to improve family planning programs.

Performance measures relate to the attainment of goals, and are generally in terms of effects, effectiveness, and efficiency. Effects, a measure of outcome or impact of program effort, can be divided into three levels, according to Reynolds (1972). The first, primary effects, deals with changes in awareness, knowledge, attitudes, and motivation. The second level measures behavioral effects (trial or adoption of a particular behavior), and the third level gauges changes in status effects (fertility, health, economic and social status).

KAP studies

The primary effects mentioned by Reynolds (1972) have mostly been measured by Knowledge, Attitude, and Practice (KAP) sample surveys. These are often undertaken in order to provide a baseline against which later studies can measure the effect of family planning programs. Much effort has been spent nationally as well as internationally in order to define terms used, as well as to standardize these surveys in order to make them comparable although programs differ in nature due to a host of cultural
factors in the situation. Nevertheless, many have voiced criticism against the KAP surveys.

Cleland (1973) claims that dissatisfaction among family planning administrators and researchers is widespread in regards to the current KAP studies. Discrepancies between survey findings and experience seem to be large, and the relevance of findings to decision-making slight. In the past, the intent of KAP studies was to measure the extent to which overt hostility to the idea and organization of family planning services existed, but today, the usefulness of these studies depends on their contribution to evaluation and guidance of existing programs. Most KAP studies have attempted to accommodate both the requirements of full demographic description, needed for international comparability and standardization, and of other needs of program administrators more directly related to operational aspects of programs. Because of practical considerations, such as interview length and analytic complexity, they have failed in this task. He recommends a future distinction between surveys whose primary objective is to provide demographic data and those focusing on operational aspects of programs.

Fawcett (1970) draws attention to the difficulty in interpreting fertility surveys and KAP studies when used for program evaluation purposes. As he views it, the dilemma stems from lack of control groups, methodological problems in initial measurement and measurement of change in attitude, and the difference in the degree of change in contraceptive practice from study to study mainly due to cultural variables.
Reynolds (1972) questions the KAP surveys because the few checks on reliability and validity that have been done on these studies have raised doubts about these characteristics. Furthermore, they do not demonstrate cause-effect relationship, they are not suited for evaluation purposes. At best, they may show correlations between program effort and KAP changes.

Because of the gap between attitude and practice, researchers have studied family planning motivational problems. Fawcett (1970) reviewed research about psychological factors in population studies conducted in the United States as well as in other countries. He groups the studies under the following headings: fertility surveys and KAP studies; studies of family interaction and fertility; persuasive communication: concepts, theories, and experiments; contraception, sterilization, and abortion; psychological factors in acceptance and use; psychological studies of attitudes, values, and motivations related to the desire for children; modernity and fertility; and, psychological consequences of family size and population density. All these psychological factors are of utmost importance when planning strategies for diffusing family planning information, and certainly need to be taken into account when explaining the success or failure of a family planning program. Nevertheless, few of these studies, and then related to poverty, are reviewed in the present study, as they are only indirectly related to the investigation.

Adoption of contraceptives

Changes in behavior have been measured largely in terms of contraceptive acceptance and continued use of these. The terms "acceptor" and
"continued use" must be qualified, as they are used to mean different things, i.e., acceptance and continued use of contraception, of a particular contraceptive method, or of a program (Reynolds, 1972; Rogers, 1973; Population Division, United Nations, 1970). Few careful follow-up studies have been conducted to determine patterns of contraceptive usage, and those undertaken rarely include reliability and validity checks. A further difficulty is to sort out the relative influence of a particular program from other intervening variables in order to determine how much credit the program can take for continued use of contraceptives (Reynolds, 1972).

Rogers (1973) points out that when a national program is initiated, rates of adoption start off fast because a receptive audience is present. Although consisting of a considerable number of people, this group has a minor demographic effect as they usually are older parents whose family size is already completed. Hardcore resisters to family planning may become adopters for spacing, rather than limiting purposes. They use contraceptives in order to attain their desired family size, and their discontinuation rate is high. Especially in less developed countries, the small family norm is not embraced by the majority of the population. To develop widespread motivation for having small families becomes the complex and difficult task for a successful family planning program (Rogers, 1973). It must build up public support for anti-natalistic policies as well as to transfer knowledge about family planning methods and to persuade their adoption.
Decline in fertility

Status changes due to the program have mostly been investigated in the light of effects of family planning programs on birth, growth, and fertility; often referred to as "birth averted". As pointed by Reynolds (1972) and Lapham and Mauldin (1972), it is exceedingly difficult to demonstrate incontrovertibly that a given program has, in fact, been responsible for a given decline in fertility. Some methods for measuring negative demographic effects are: changes in fertility status of one group over time; comparison of actual versus expected changes in fertility status of one group; changes in fertility trends in experimental groups (versus control or matched group, or total population); and indirect calculations of program effects on fertility, based on estimates of "couple years of effective contraception that can be attributed to contraceptives distributed by the program" (Reynolds, 1972, p. 76).

Program efficiency

Program efficiency is often expressed in ratios, such as births averted per dollar, patients seen per physician hour, and then compared to a goal, standard, or competing program.

The input/output, such as cost per acceptor, is a measure of efficiency that has been found useful. To estimate the cost/benefit or effect, is more complicated, mostly because of the difficulties in defining and measuring behavior and status changes.

Theoretical issues

Many theoretical issues are involved in the evaluation of family planning programs at the national level.
Population programs is a broader concept than the delivery of family planning and abortion services. They may include such aspects as health care, population education, legal statues, and social trends like the rising age at marriage. Nevertheless, voluntary family planning services remain the leading, and in many countries the only prominent type of programs designed to slow rates of population growth, while concurrently bringing about improved maternal and child health (Lapham and Mauldin, 1972).

Another conceptual problem is the controversy of the idea of family planning. Reynolds (1972) submits there are at least three types of family planning programs: those that admit they want to bring down birth rates; those who deny they want to do so; and those who do not want to say. The first type can be evaluated in terms of their success in reducing fertility, but programs that emphasize voluntarism, birth spacing, or having the number of children desired, should not be measured by this criteria alone.

Most evaluations take the family planning program as the independent causal variable of fertility decline. Reynolds (1972) proposes that, instead, it is more likely that family planning programs are one of many intervening variables that contribute to fertility regulation under certain conditions. Hence it might be more appropriate to study the conditions under which family planning programs contribute to fertility regulations.

A further debatable point is the value of setting specific program targets, as it on one hand can increase motivation and achievement, but
also lead to falsification of records at various levels, coercion of adopters, etc. Another problem is how to set targets, at what level, over what period of time, etc. (Lapham and Mauldin, 1972; Rogers, 1973).

Many technical problems arise in the assessment of family planning programs. Reynolds (1972) lists: lack of indicies and scales for the measurement of health, motivation, contraceptive behavior, and other key variables; lack of analytic techniques for measuring small changes over short period of time; and lack of attention to nonprogram, intercultural, and intracultural variations. Difficulties involved in fitting family planning programs into an overall framework of economic, social, and planned development provide a hardship for evaluating the success of these programs.

**Systems for evaluating national programs**

Several authors suggest systems and frameworks for the assessment of national family planning programs. As these have many characteristics in common, only two are reviewed.

Bogue (1970) argues that an evaluation program, in order to be truly effective, must be nationwide in scope. He recommends an evaluation system consisting of four components. One of these would be demographic analysis, dealing with birth rates and target population and quotas; another national family planning sample inventory every three years as to motives, attitudes, knowledge, explanatory variables, use of contraception and pregnancy prevalence. A third component is labeled national evaluation of family planning services, studying clinic reports on new patients, follow-up services, personnel, expenditure, and with follow-up studies of clinic
patients every three years. Lastly, special studies by individual persons and institutions of particular projects that do not fall into the other three categories would compose a fourth component.

Lapham and Mauldin (1972) propose a framework of evaluation that consists of four sections, namely (1) statistical measurements concerning acceptors and users; (2) procedural and supportive measures that have been adopted by some programs that may have some relationship to increased acceptance and use of contraceptives and maternal health services; (3) the fertility reduction measurements themselves, that is, declines in age specific fertility rates and crude birth rates, and the degree to which such declines are attributable to the programs; and (4) social and economic phenomena that may be related to fertility reduction processes. Social, health and mortality, economical, historical, cultural, religious, and demographic matters are difficult to isolate as factors. Nevertheless, it is necessary to review the social and economic setting within which the program operates in order to determine to what extent they create conditions favorable or unfavorable to the program. Review of evaluation studies

National level An attempt to evaluate the effects of population control programs, defined as the deliberate effort by governments to affect the rates of population growth in their countries, was made by Berelson (1974). In his theoretical framework, he states that there are five things governments can do. One is to communicate with people through factual information, enlightenment, or persuasion, in order to influence their demographic behavior in the desired manner. A second thing is to
provide services to effect the desired behavior. Thirdly, governments can manipulate the balance of incentives and disincentives to achieve the desired regulation, i.e., raise or lower the cost of having children. A fourth way is to shift the weight of social institutions and opportunities in the desired direction, i.e., regulation of marital status, age at marriage. Lastly, people can be coerced to the desired behavior through the power of the state.

Urbanization and industrialization with their impact on kinship and belief systems, popular education, nutrition and sanitation, and liberation of women are elements of social change that, according to Berelson (1974), provide the motivation for smaller families by increasing the economic cost of children, by undermining the sense of fatalism, and by providing the parents, and particularly the mother, with alternate opportunities of so-called lifestyle.

Berelson (1974) cross-classified 26 developing countries with national family planning programs according to their family planning programs. The countries were classified by their relative development on three major indicators, presumably associated with motivation for a small family: per capita gross domestic product as an indicator of standard of living; infant mortality as an indicator of health, and female enrollment in formal schooling as an indicator of popular education. By dividing each ranking into thirds and then constructing a composite index, each country was categorized overall as high, middle, or low in motivation.

Independently, the family planning programs in these 26 countries were classified as strong, moderate, or weak, based on their coverage,
their continuity and duration of effort, the vigor with which the program was actually pursued, and so on.

Both factors are important as at each level of program strength, the developmental setting makes about a threefold difference in annual acceptance of contraception, and at each level of development the programmatic effort makes a twofold difference.

Program effort makes more of a difference where the setting is at least moderately encouraging, and seems somewhat more powerful in cumulation. Berelson (1974) warns for any judgments that family planning is either impossible in the absence of modernization or unnecessary in its presence. He notes that there is an implicit relationship between setting and program: the more modernized countries can carry out stronger programs, and the less modernized ones can build up even a moderate program only with considerable effort. Thus, the setting conditions not only response but programmatic effort as well.

The causal analysis of fertility changes is extremely difficult to carry out with definite results. Declines in birth rates attributable to programs were only claimed for four countries in the analysis. In the other instances, the decline was too small to show itself, or could not be ascribed to the program on technical grounds. The four countries had a low death rate, even at Western standards. This Berelson (1974) considers a threshold factor of importance to fertility.

If early stabilization of population is the goal, the present type of family planning programs will not reach that goal, but if actuated it will move in that direction and reinforce and expedite an inevitable process of
fertility control. It seemed to Berelson (1974) that, today, governmental policies to reduce fertility rates were proving more effective than governmental policies to increase fertility rates.

What governments should do to control population growth through fertility, according to Berelson (1974), is three things: maximize the flow of full information about the consequences of fertility behavior, both individual and collective; maximize the capacity of people to regulate their fertility by medically approved means, in accordance with individual conscience; and minimize undue pressures on fertility behavior either way. Where these conditions are met today, the fertility is very low.

**Local level** When evaluating the effectiveness of local family planning services, Jackson (1972) advocates that the congruence or dissonance between behavior and attitudes towards fertility control be examined. Behavior in this context is defined as the dissemination or diffusion of birth control information to others. Jackson (1972) contends that the general effectiveness of a voluntary birth control effort is linked to the diffusion of favorable comments in informal discussion between relatives and friends in the neighborhood environment. Willingness to convey information is dependent on knowledge of and approval of family planning clinics. Variables influencing the knowledge of birth control clinics are the proximity to the clinic, the friendliness of the neighborhood, and the person's propensity to discuss birth control. According to Jackson (1972), variables that influence birth control clinic approval are the respondent's age, education, and family size; fundamentalism and
fertility, and other attitudinal determinants such as belief in the efficiency and safety of birth control techniques and the conviction that some families would be better off with fewer children. Willingness to convey birth control information to others is dependent upon belief in the efficiency or effectiveness of birth control techniques in controlling family size, belief in the safety of the method in terms of woman's health, and belief that some families would be better off with fewer children. In order to test his hypotheses, Jackson (1972) conducted a study among low-income families in a metropolitan area. Among the findings were that educational attainment is conducive to lower birth rate. For many women in the study religious conviction superceded other considerations in decisions concerning the number of children a family will have. There was an apparent strong relationship between attitudes toward the safety of contraceptives and willingness to diffuse information concerning the use of these contraceptives.

Jackson (1972) proposes that family planning programs emphasize health aspects of birth control devices and that efforts be made to locate the clinical resources as close to prospective users as possible, in order to increase awareness of and responsiveness to the program.

Family planning and low-income audiences

In the United States most all people have come to share a quite similar set of fertility values and practices (Fawcett, 1970; Jaffe, 1964). A majority of middle-class and upper-class North Americans had adopted contraceptive methods by the mid-1960's (Rogers, 1973). However, only a small portion of low-income people were using family planning
methods, not always because of lack of motivation, but because of the nonavailability and nonaccessibility of good clinic services (Jaffe, 1964; Rogers, 1973). Recently, by law, the federal government requires each state to make family planning services available to low-income people.

Rainwater (1960, 1965), using intensive interview techniques, studied conjugal role relationships and patterns of sexual behavior in lower and middle class urban and recently urbanized Americans and related these variables to family size preferences and contraceptive practices. In regards to educational efforts made through direct and massmedial techniques to inform and persuade about contraception, Rainwater (1960) found that within the working-class there is a market and an audience for material about sex and contraception. As most of the respondents in his study knew about two and more contraceptive methods, he suggested that the problem is not one of simple knowledge but of making knowledge meaningful to the people themselves. Many working-class people have little understanding about their own and their spouse's genitals and about contraception and have difficulties in understanding the difference contraceptive methods make. Rainwater (1960) proposed that if educational efforts enable the people to link the method itself with an understanding of what it accomplishes, the adoption of effective contraception might be more likely. Although articles in massmedia, over a period of time may contribute to a change of attitudes and prepare the grounds for contacts with clinics and personal physicians, Rainwater (1960) cautions that many of these articles are confusing to working-class readers because of their general ignorance in these
areas. A book, rather than a comic magazine, explaining sex and contraceptive matters in a simple and meaningful fashion, perhaps using the approach of the better "family problem" columnists, might find a reading audience among working class people. Rainwater (1960) further states that well-conceived advertising campaigns might be another source of education, because this is a method through which most Americans learn about technical innovations. Especially the working-class learn to regard as desirable many activities that require accessory products through advertising.

Hochstrasser et al. (1973) noted a fertility decline in Southern Appalachia during the last four decades, the most precipitous decline occurring in the 1960's and 1970's. Out-migration of certain age groups could only explain about 20 percent of this trend. Research conducted during 1964-1967 showed a level of awareness and approval of birth control and family planning practices approximate to relatively high national norms at that time, but also a significant lower level of specific knowledge and actual usage of contraceptive measures and family planning practices. Thus fertility was declining despite the fact that the general use of contraceptives remained low. Hochstrasser et al. (1973) found support from other studies conducted at the same time for this finding of comparatively high level of acceptance coupled with low level of adoption of birth control and contraceptive practices. Drawing a parallel with the fertility decline during the depression years in the 1930's, Hochstrasser et al. (1973) suggest that possible explanations are in as yet unidentified social and situational factors related to the response of individuals and families to the conditions of relatively severe economic and social deprivation and overall poverty.
One of the few longitudinal studies made of poverty and family planning was conducted in Finland, where Visuri (1969) investigated family planning practices of couples receiving temporary and regular public assistance in Helsinki. A total of 100 families were followed through public welfare case records from 1956 through 1965, when the couples were interviewed.

By comparing families receiving temporary and regular assistance, Visuri (1969) hoped to highlight the intensified problems of family growth. Both groups had approximately equal numbers of children when they first needed assistance, but later on, families receiving long-term assistance had on the average one child more than those on limited assistance. Family growth in the first group most often occurred simultaneously with the need for assistance. Although family growth had been experienced as a problem in the regularly assisted families, birth control was not always practiced as a solution to this problem.

The study used the opinion of the respondents as to success with their contraceptive method, number of unwanted children, and number of unwanted pregnancies as measures of success or failure in the use of family planning methods. The level of information on contraception when the wife first did not want a child was found to be the most important antecedent factor influencing the number of unwanted children and pregnancies. Wives with more traditional reference groups and from lower socio-economic strata were less informed on contraception. The "time factor" might have influenced this finding as family planning organizations and methods were not so readily available in the earlier years of the study.
In 1965, every mother had knowledge of at least one contraceptive method, and a large majority knew about all the methods asked about, but the level of information had been lower during the earlier years of marriage. The "best" method by which the couples were most often successful according to their own statements, was the withdrawal method, which for the majority was known from the beginning of the marriage. Visuri (1969) therefore suggests that the number of methods known may only be an indication of general enlightenment.

The reasons for not wanting a child were in half of the cases financial, according to the respondents in the study, and economic criteria were most often used in justifying an ideal number of children.

Visuri (1969) concludes that still today many families in Finland have a large number of children because of unwanted fertility, and some of them are poor. But there are also couples who would like to have a second or third child, but who in order to maintain even a modest standard of living have to abstain from it. This finding can be compared with Rainwater's (1965) conclusions that the rationale for large and small families among middle-class and working-class people in the United States is that one should not have more children than one can support, but one should have as many children as one can.

Few studies have been found that relate directly to the present study. Although Brand (1971) investigated the status of, and need for, family planning information and communication in the Expanded Food and Nutrition Education Program (EFNEP), as perceived by the paraprofessional aides, only one investigation is known to the author that attempts to
evaluate a program intending to diffuse family planning education. In this research, Holt (1975) evaluated a two-week in-service seminar with home economists in Panama. The aim of the study was to assess factors leading to the actual utilization of family planning education, an innovation introduced through the seminar.

Criterion for effectiveness of the in-service program was praxis, viewed as a combination of reflection and action. Praxis was measured by the number and type of family planning education activities actually planned and carried out by participants after the seminar, as well as by the number of related concepts they reported having taught through their program of work.

Commitment, measured by the number and type of individual goals set by seminar participants for themselves at the end of their in-service training, was hypothesized to be an intervening variable, contributing directly to praxis.

It was further hypothesized that seven independent variables contributed both to praxis and commitment: knowledge of family planning; experience with family planning education; involvement in school and community affairs; leadership demonstrated during seminar group discussions; perception of support of significant others for participants' involvement in family planning education activities; favorable and liberal attitude toward family planning; and internal locus of control of reinforcement.

Data were gathered through pre-seminar, post-seminar, and follow-up questionnaires, and through observations by group discussion leaders. The self-administered praxis measurement was validated with interviews.
The pre- and posttests involved 176 home economists, of whom 87 or 52 percent, were included in the follow-up study.

Analysis by correlations and multiple regression support the use of the hypothetical model in the study at the .01 level. Although only 25 percent of the total variance was accounted for by the combination of the variables in the study, all contributed to the prediction of praxis. Commitment contributed the greatest amount, and was in turn predicted by the remaining variables at the .05 level of significance.

Holt (1975) found a strong relationship between praxis and commitment. Commitment is generally viewed as a key factor in the adoption of an innovation, but in order to maintain its strength, it must be the result of uncoerced decisions made by the individuals, or goals they set for themselves. Public pronouncement of an individual's intentions might result in stronger commitment than a privately made one.

While all variables contributed to the prediction of both praxis and commitment, depending on the value of the variable, control, they proved to vary in importance. Internally controlled individuals became committed to a greater extent in the study, but the externally controlled, once committed, followed through more faithfully with their commitments. Perception of support of significant others seemed to be important for externals. For the internals, their own attitudes toward family planning were important.

As leadership and internality were strongly related, Holt (1975) suggests that leadership training might be a useful tool for indirectly increasing commitment.
At the time of the post-seminar evaluation, positive attitude toward the innovation was correlated only with knowledge, but in the follow-up study positive attitude toward family planning in general, and toward the role of the home economist in family planning in particular, correlated significantly with praxis. Knowledge of the innovation was at best a minor factor in the eventual infusion of the innovation into programs.

Applying different program approaches to internally and externally controlled individuals; working toward a change in locus of control; concentrating on strengthening one variable, such as leadership, were suggested by Holt (1975) for making in-service training more effective. In addition, mass media campaigns, support of administrative levels, and backing of religious leaders were all favorable factors in the success of the experience in Panama.

Summary

This section has traced the development of the evaluation of educational programs. The major trends within this area have been explored, and the theoretical framework of the most important models discussed. An attempt was made to demonstrate the various ways in which these models have been applied in practice. Other literature cited has dealt with program evaluation within the Cooperative Extension Service. Special attention has been given to programs aimed at low-income adults. Lastly, the literature was searched for evaluation efforts of family planning programs.
Based on the review, it appears that in order to optimize an educational program, evaluation needs to be an integral part of program planning and execution. Any educational program requires constant adjustment due to unremitting changes in the social, economic, and intellectual needs of the society. In a pilot program, where many decisions have to be made without support of prior evidence, i.e., with a high risk-taking, continuous assessment may reduce the uncertainty on the part of the decision maker.

The Context, Input, Process, and Product (CIPP) model (PDK Study Committee, 1971) offers a holistic, directive, broad base for evaluation. It provides a means through which it is possible to identify key decision points in the program, and to predict more easily the consequences of a decision at a particular moment. Furthermore, it becomes feasible to anticipate or detect specific weak or strong links within the program and hence strategies or revisions can be planned and put into effect at an early stage so as to maintain desired outcomes.

The literature also points to the fact that experimental studies are not well suited to either social action programs or to pilot programs, even if for inferential purposes this method needs to be approximated as closely as circumstances permit. Moreover, a program evaluator needs to collect a wide range of evidence, and of intended as well as of unintended effects of a program. Much of the data, especially in the case of programs geared towards low-income and/or ethnic minority groups, will be of unobtrusive, nonoffending nature, thus falling into the category of descriptive data.
METHODOLOGY

Project Under Study

A pilot project, aiming at including family planning education within the established Expanded Nutrition (ENP) program received special funding within the Cooperative Extension Service in Iowa. As it was deemed important to evaluate the introduction of the new concept in order to facilitate future decisions regarding the continuation, expansion and/or termination of the pilot project, the Department of Home Economics Education in collaboration with extension undertook to assess the program. The present investigator served as an internal evaluator for the initial part of the project, viz. March 1974 through June, 1975. At this point in time, the pilot project had gone through one cycle of development. It had been initiated, a state planning committee was operating, local extension home economists were involved, and aides had been trained to relate family planning concepts to decision-making as well as to the different areas of home economics. They were actually working in the field, and using material developed for this purpose.

Theoretical Framework

The theoretical framework for the study was a modification of the Context, Input, Process, Product (CIPP) model. The notion of assessing a program in relation to the context in which it operates, the input of human and nonhuman resources, the process, and the product in terms of intended and unintended outcomes provided a holistic and directive, broad
base for evaluation. For the project under study, the trial stage of the CIPP model was deemed appropriate. The decision setting for the project was classified as incremental. The intended change was to introduce the concept of family planning into the already established teaching scope of the paraprofessional aides in the Expanded Nutrition Program (ENP). As discussed in the previous section, this decision setting is characterized by low information grasp and a low degree of change. Using expert judgment and structured inquiry, a problem analysis is done and apparent change possibilities tentatively selected and tried out until a successive approximation of a solution is reached.

Due to a variety of circumstances, the author was not able to completely follow the PDK Committee's (1971) methodology, i.e., delineate, obtain and provide useful information for judging decision alternatives. Features of both the goal-free and the responsive evaluation were included, such as collecting evidence as broadly and extensively as possible from the goal-free model, and the evaluator functioning as a sounding board for the program manager, giving quick, personal judgments about the program from the responsive evaluation model.

Objectives of the study were:
- to record and describe the planning and implementation of a pilot program;
- to analyze the processes and procedures used in the program; and
- to make recommendations about processes and procedures, based on the analysis for other similar programs.
Assumptions underlying the study were that subjects involved in the program planning and execution would be willing and able to provide needed information; that the CIPP model could be entered at any state; and that information gained from the pilot program could be useful in other similar programs.

Many limitations surround the present study, all of which influence the findings. It was limited to one pilot program. The number of subjects involved was small, making statistical analysis difficult and hence the study descriptive in nature, and making inferences outside the specific program very tentative. Because of the early stage of the program and the limited time the author could participate in the pilot program, major emphasis was on the process evaluation. Also, the investigator had to rely on periodic written and oral communications instead of day-by-day involvement in the project. As the author was not involved in the program until March, 1974, the context and input evaluations had to be done in retrospect rather than precede the planning decisions, as the Context, Input, Process and Product (CIPP) model advocates. Because the data-gathering ended before the program was firmly established, it was too early to study the effect of the program, hence product evaluation is minimized in this study.

Participants in the Study

Although the population in the study was defined as all persons involved in the pilot program planning and execution, i.e., administrative personnel and subject matter specialists at state and local levels, para-
professional personnel, and program participants, only those directly responsible for the program did participate in the investigation.

It is recognized that formal and informal legitimizers, such as extension area director, local home economics advisory council, other local agencies, and state extension administrators are of importance when establishing a new program, but in order to limit the scope of the study, and because of the early stage of the program, these were not contacted directly by the researcher.

Participants in the study were the state leader of the home economics programs, the assistant state leader directly in charge of the ENP program and the pilot project, six state extension subject matter specialists, five local extension home economists and the paraprofessionals working on the project. The six state home economists were specialists in food and nutrition, health education, housing and home management, family life and human relations, child development, and textiles and clothing. Of the local home economists, one was in charge of the ENP and the pilot project, one was the area specialist in family life and human relations; the others had their areas of concentration in food and nutrition, housing and home management, and housing and interior design. The number of aides involved varied during the study. At the beginning of the training of the paraprofessionals, all those employed in the ENP program were invited to attend. Final assignment of eight half-time aides to the project was made in August, 1974. Due to turn-over among the aides, some fluctuation in the number working in the project at any one time occurred.
Formulation of Questions to be Answered by the Study

In order to structure the study, questions that might be answered by the research were formulated by the investigator after a search of the literature and informal discussions with extension personnel. These questions were later used as a guideline when deciding on what data to collect.

As the CIPP model was used, the questions were grouped under the headings of Context, Input, Process and Product. A broad, overall question was posed under each heading, followed by specific, more detailed ones. Thus a total of four major and 26 minor questions were asked, namely six minor ones under Context, five under Input, 13 under Process, and two under Product. A list of these questions can be found in the Appendix.

Because the pilot project operated in the trial stage of the CIPP model and because the product was only starting to emerge owing to the early stage of the project, major stress in this study was on the process. This practice is recommended by the PDK Committee (1971) and Macy (1975) for the early stage of a trial. Nevertheless, as process in isolation cannot be fully understood, context, input and product were also considered, although to a lesser extent and mainly in the instances they did deviate from the regular ENP program.

The questions were submitted to a total of five persons in order to select ten which were considered most important for future decisions about the program. The five persons included an associate and an assistant dean of the College of Home Economics, one of whom had been director for inter-
national family planning workshops, and the other of whom was the state leader for the home economics program in extension; one professor in the Department of Family Environment, specialist in family life and co-director for an international family planning workshop; and two professors in the Department of Home Economics Education, both with extensive experience in extension work and familiar with the ENP program, one specializing in evaluation and the other in adult education.

A tally was made of the votes for important questions to be answered by the study. Consensus was reached on one question only, i.e., number 1 under Context, and number 4 under Process. Four questions received four votes each, eight three votes, five two votes and six one vote, respectively.

The final list of questions was made up by those receiving the most votes. In a few instances, it was deemed possible to combine two questions into one. The list emerged as follows:

**Context**

In what context were the goals and objectives of the pilot program formulated?

1. What theoretical information was collected before the decision on goals and broad objectives of the pilot program were formulated?
2. What information (evidence of needs, study of community, existing services) was collected before decisions on goals and broad objectives of the pilot program were formulated?

**Input**

What resources were available, and how were they to be utilized in order to meet program goals and objectives?
1. What human and nonhuman resources and obstacles within and outside the organization were identified during the planning stage?

2. What strategies were selected on what basis for achieving the program goals?

Process

What were the main features of the project design, and which ones contributed to the success or failure of the design to reach program goals and objectives?

1. What changes were made during the implementation - based on what information, what alternatives, using what decision rules?

2. What kind of leadership was provided? Expected?

3. Did people involved in the program understand and agree with the intent of the project?

4. What training methods were successful with the paraprofessionals - as viewed by the paraprofessionals, their supervisor and the trainers?

5. What training was provided for the local home economists? For the state specialists?

6. What approaches did the paraprofessionals use that they consider successful or failures?

Product

To what extent were the program goals and objectives attained?

1. What were the accomplishments of the pilot project as related to the program's short-term and long-term goals?

2. What kind of information was needed for recycling decisions?
The questions were used as a guideline for data-gathering. Due to the pilot nature of the project and circumstances beyond the control of the investigator and sometimes beyond that of the state and local program managers, it was not always possible to collect data to answer questions in depth. This will be discussed in the Findings and Discussion section.

Context

The theoretical context in which a program operates includes the general trends in the contemporary society and the basic philosophy of the sponsoring organization and its members. For the program under review, ideologies about education, man and decision-making, socio-economic and ethnic minorities, overpopulation and its effect on society and individuals, were of special interest.

Evidence of needs for including family planning education within the teaching scope of the ENP aides might include demographic data about program clientele, request for family planning information from the clients, perceived need by other agencies working with the same or similar clientele. A general demographic study of the community and of existing services in the area of family planning might testify to the need for additional service in this area.

Input

For any program planning, it is important to investigate what resources are available, not only in form of funds but also such things as physical facilities, trained manpower within and outside the organization, positive attitudes toward the program content in the community, the
organization and individuals, availability of trainees and possible audience.

Innovations are subject to resistance from community, organizations, and individuals. It is valuable to attempt to establish the particular obstacles the innovation may encounter. If known, they can be counteracted in the strategy for diffusing the innovation. Possible obstacles for family planning education include such things as cultural values and personal attitudes of clients, paraprofessionals, professionals, and in the community-at-large. The relative taboo-nature of the communication, negative rumors about the subject matter, contradictory messages by other persons or organizations may well work against the intended educational program.

Based on information gathered about context, available resources, and possible obstacles it is possible for a program manager to select a strategy for program implementation. Usually more than one solution is possible, but one judged to be superior. It is a value judgment on the part of the decision-maker and from an evaluation point of view it is important to know not only which strategy was chosen but also the criteria for the choice.

Process

A program is apt to change during the implementation stage as circumstances alter. This is particularly true of a pilot project unless it follows the experimental model. In order to assess a program, information about changes that occurred and underlying reasons is needed. Knowledge
about perceived alternative actions would create a fuller understanding of the program.

Leadership is important in any undertaking. Leadership styles vary and no one can be categorically deemed as best in all situations. Incongruence between provided and expected leadership on the other hand, may create an obstacle in program execution as it can lead to inactivity, delays, or even resistance among program personnel. It is the perception of the individual that is of interest as she/he reacts to the leadership as she/he perceives it.

In order to create and maintain positive support and action on the part of the program personnel, it is vital that the intent of the program is clearly understood and agreed upon. In the opposite case, resistance may be built up and the smooth implementation of the program impaired. It is the individual person that might be tapped for this information.

What constitutes a successful method of training paraprofessionals has been the topic of several research studies. Nevertheless, it is important to find methods that all levels of personnel are comfortable with and yet successful in transmitting the intended message in an understandable and acceptable form, as well as that motivate the aides to include it in their teaching. Personal opinion, formal tests, logs kept by aides, and questions brought up by aides might be sources of information in this respect.

As family planning education is a new teaching concept for all levels of extension personnel, training in this subject matter may be appropriate. Due to the nature of the topic, a certain familiarity with the
subject matter as well as an opportunity to practice talking about it might counteract the inhibition otherwise felt by a person when coping with it. Information about training can be found in state and local files as well as collected through observation during execution.

The paraprofessional aides are the ultimate link with the families. Although an approach to family planning education may well vary from aide to aide as well as from family to family, it might be possible to find commonalities in what the aides perceive as successful or unsuccessful approaches. The aides were the source for this information.

Product

Programs usually have both short-term and long-term goals, although often one of these is implicit rather than openly stated. In a pilot project the long-term goals are rarely visible at the time of its assessment, hence only short-term goals lend themselves for evaluation. In the program under review, evidence of progress towards objectives was sought through personal interviews with program personnel, evaluation undertakings by the project managers during the project and information available in state and local files.

Information needed for recycling decisions was determined through interviews with decision-makers.

Data Collection

Data were collected from March, 1974 through June, 1975 through unstructured observation of state planning committee meetings, informal discussions with the assistant state leader, formal and informal inter-
views with six state subject matter specialists and five local home economists, as well as through unstructured observation of training meetings of local home economists and paraprofessional aides. Comments from the aides about their training, results from internal efforts to evaluate aspects of the project, and information about the project available in files at the state and local levels have been utilized in this study.

Eleven state planning committee meetings were held during the period of the study, mostly on a monthly basis. During 1974 the meetings were in March, April, June, July, August, September, November and December. In 1975, meetings were held in March, April and May. During the meetings, unstructured observations were made about content, interaction, decisions made, and overt reaction of participants.

Informal discussions were held with the assistant state leader throughout the study. The frequency of these depended on the particular pace of the project. Notes were kept about content of these discussions.

A formal interview was conducted with each of six state subject matter specialists and with each of five local home economists in August and September 1974. At this stage only one specialist had conducted a training meeting at the local level. The interview concentrated on the interviewees' knowledge and experience with low-income groups and in working with other ethnic groups, their perception of integrating family planning with their subject matter area, as well as of their's and others' role in the program. Their opinion about the ultimate outcome of the project was solicited as was their perception of who were the decision-makers in the project. Only a slight alteration in the questions was made
for the two groups. Informal interviews were held whenever opportunity occurred, especially after training meetings, when training methods and the reaction of the audience(s) were discussed.

Training meetings for local staff and paraprofessionals were held in 1974 during August, September and December, and during 1975 in February and March. Unstructured observations were made of content and of method, as well as of the participants' interaction and overt reaction to the training. The ENP home economist was interviewed about the aides' reactions to the training sessions as the aides felt free to discuss these with her on later occasions. The home economist kept some records of these comments, which were made available to the investigator.

In order to facilitate planning and to have immediate feedback, the assistant state leader endeavored to evaluate aspects of the program as it developed. Thus an attitude and knowledge questionnaire was produced and administered locally to the aides in May, 1974, as was another questionnaire in April, 1975, soliciting information about their understanding of the project, of the relationship between family planning education and the different areas of home economics, their perception of how to improve their training, and of sources of frustration in their teaching efforts. At the same time, members of the family planning committee were asked to evaluate their participation in the project. The results of these questionnaires were made available to the investigator.

The development of instruments or recording forms to collect information from and about the clients by the aides was followed by the investi-
gator and she had some opportunity for input into these. The same was
the case with visual aids to be used by the aides in their work.

Project information available in files at the state and local level
was made accessible to the investigator without restriction. This in-
cluded the logs written by the aides after each individual visit, routine
demographic information about program participants and the aides, and
administrative correspondence.

Treatment of Data

The findings are presented as responses to the final twelve questions
about the project, the formulation of which was described earlier in this
section. These findings were derived by synthesizing data collected from
observations of training sessions and planning committee meetings, from
interviews with personnel, from official files, and from internal evalua-
tions by project staff.
FINDINGS AND DISCUSSION

The program in this study was a pilot project within the Expanded Nutrition Program (ENP) of the Iowa Cooperative Extension Service. The innovative part of the project was to broaden the teaching scope of the paraprofessional aides to include family decision-making as it relates to family planning, and to integrate this with concepts in home economics subject matter areas: nutrition, health, human development, housing, textiles and clothing.

This study encompasses the initial period of the project with data collected between March, 1974 and June, 1975. The objectives of the study were:

- to record and describe the planning and implementation of a pilot program;
- to analyze the processes and procedures used in the program; and
- to make recommendations about processes and procedures, based on the analysis, for other similar programs.

Questions to which the study speaks, were developed as described in the Methodology section. The final list emerged as follows:

Context

In what context were the goals and objectives of the pilot program formulated?

1. What theoretical information was collected before the decision on goals and broad objectives of the pilot program were formulated?
2. What information (evidence of needs, study of community, existing services) was collected before decisions on goals and broad objectives of the pilot program were formulated?

Input

What resources were available, and how were they to be utilized in order to meet program goals and objectives?

1. What human and nonhuman resources and obstacles within and outside the organization were identified during the planning stage?

2. What strategies were selected on what basis for achieving the program goals?

Process

What were the main features of the project design, and which ones contributed to the success or failure of the design to reach program goals and objectives?

1. What changes were made during the implementation - based on what information, what alternatives, using what decision rules?

2. What kind of leadership was provided? Expected?

3. Did people involved in the program understand and agree with the intent of the project?

4. What training methods were successful with the paraprofessionals - as viewed by the paraprofessionals, their supervisor and the trainers?

5. What training was provided for the local home economists? For the state specialists?

6. What approaches did the paraprofessionals use that they consider successful or failures?
Product

To what extent were the program goals and objectives attained?
1. What were the accomplishments of the pilot project as related to the program's short-term and long-term goals?
2. What kind of information was needed for recycling decisions?

Data were collected from March, 1974 through June, 1975 by means of unstructured observations of eleven state planning committee meetings and of six training meetings for program aides and local home economists. Formal and informal interviews with state home economics specialists and local home economists and informal discussions with the assistant state leader were held. Results from internal efforts to evaluate aspects of the project, and information about the project available in files at state and local levels were additional sources of information.

Throughout the study, the investigator served as a sounding board to the assistant state leader and provided quick, personal judgments about the different aspects of the project. The judgments were based on observations of the project, discussions with persons involved in the project, theoretical information, and personal experience. It must be kept in mind at all times that the nature of the pilot project was a developing one, rather than a purely experimental one. Thus, although objectives and strategies were outlined in the proposal of the project, they were subject to change as circumstances varied.

A developmental pilot program is apt to change as conditions change. Repeated change makes it difficult to separate out useful components of a program from those that are less so, or from those that are counterpro-
ductive. Suchman (1970) suggests that in the pilot stage of a program, in which new approaches and new organizational structures or procedures can be tried out on a flexible basis, a quick-and-easy type of evaluation with primary emphasis upon feedback for program changes be employed. The same thought is behind the evaluation of broad-aimed programs, where the issue is "What happened?" rather than "Does it work?" (Weiss & Rein, 1970) and in goal-free evaluation (Scriven, 1972) where the concern is to determine the overall effects of a program, whether intended or not. Weiss (1972a) maintains that evaluation can be premature in cases where the assessment begins before the program has found its goals, its function, or generally accepted ways of work. In these instances she suggests that an analysis of the system, rather than evaluation will be helpful to the program developers.

Therefore, program evaluation that establishes cause-effect relationship between the program and occurred changes will not be possible until a program has established itself. A benchmark study of existing conditions done in the early stage of the program development would immediately be valuable for program planning purposes and in the longer term for evaluation purposes.

Context Evaluation

According to the PDK Committee (1971, p. 218), the purpose of context evaluation is to provide a rationale for determining objectives. It defines the environment, describes the desired and actual conditions pertaining to that environment, identifies unmet needs and unused oppor-
turnties, and diagnoses the problems that prevent needs from being met and opportunities from being used.

Context evaluation in the CIPP model is fundamentally philosophical, in that it describes the values and goals of the system under investigation. At the same time, it also reflects the theoretical and empirical knowledge in the field and seeks to determine whether practice is consistent with the validated principles of a relevant theory (PDK Committee, 1971).

As the investigator was not involved in the project at this stage, information was collected retrospectively through recall interviews and through filed materials. Planning decisions were made at the higher administrative level.

In the study, one main and two specific questions were stated under Context evaluation, viz.:

In what context were the goals and objectives of the pilot program formulated?
1. What theoretical information was collected before the decision on goals and broad objectives of the pilot program were formulated?
2. What information (evidence of needs, study of community, existing services) was collected before decisions on goals and broad objectives of the pilot program were formulated?

Theoretical information

Theoretical information influencing the pilot project included concern about population and its effect on the living standards, especially
on a low-income level, and the philosophies of the American Home Economics Association (AHEA) and the Cooperative Extension Service.

Family planning has gained increased attention at world, national, and local levels as a result of a mounting awareness of the rapid population increase and its effect on the limited resources on earth. A zero population growth (ZPG) rate, or replacement level only, becomes more and more an accepted and stated goal of nations, as too rapid a population growth may seriously hamper efforts to provide sufficient food, health care, housing, education, as well as cultural and recreational opportunities for the expanded population.

A study of fertility behavior in the United States in 1965 (Ryder & Westoff, 1971) indicated that 19 percent of births were unwanted. The proportion of unwanted births increased among respondents with larger families and for those with less education. Between 1966-1970 the figure for all unwanted births had changed to 15 percent, but the trend remained that the number was higher for certain strata in society (Presidential Commission on Population Growth and the American Future, 1972). To eliminate unwanted births in the total population would not in itself ensure the ZPG goal be reached in the U.S., as the number of wanted children still exceeds the replacement level, although the gap is rapidly narrowing.

The cause-effect relationship between poverty and large families has not yet been established, but as each child in a large, poor family gets a poorer start in life than if the family is small, this is likely to be one way to extend poverty into the next generation. Although the low-income
families have proportionately more unwanted births, they represent a small portion of American society relative to those with higher incomes, and therefore preventing their unwanted births might contribute little to lowering completed fertility in the entire society (Bauman, 1972). On the other side, the effect of unwanted children in low-income families is more drastic than in families in other socio-economic levels. Low-income women are usually less educated, hazy about reproductive physiology, and ill-equipped to select an effective contraceptive (Oppenheimer, 1973). Furthermore, this category of women does not have adequate access to competent and supportive advice from social workers and doctors, particularly if they are not married, even though a recent federal law requires all states to provide family planning services to any person who needs them but cannot afford to obtain them through his or her private physician.

The AHEA (1972) endorsed a declaration by the United Nations which states that couples have a right to choose freely and responsibly the number and spacing of their children. It also subscribes to the idea that children have a right to be born into families where they are wanted and loved, as well as to the idea that planning for size of the family is one kind of planning that families can do to improve family living.

The association furthermore envisages that home economics can help combat the population problem through assisting family members in realizing the relationships between family goals, family resources, and family size. This might be achieved through an understanding of how a rapid population growth rate affects living conditions for families, and
of family planning as a decision-making process related to rights of parents and children.

The pilot project under investigation operated within the framework of the Expanded Nutrition Program (ENP), an integral part of the Iowa Cooperative Extension Service. Thus many features, such as philosophy, administrative organization, facilities, staff, and the use of paraprofessional aides, were shared with extension in general and with the ENP program in particular.

From its beginning, the Cooperative Extension Service has been conceived as an educational program, through which knowledge could be transmitted from the researcher directly to the people. It focuses on the individual and the family, enhancing human development and maximizing the individual's contribution to society. The extension educator is viewed as a change agent who creates learning situations and serves as a resource person to his/her clients in problem-solving (Wang, 1974). Traditionally, extension programs have centered around agriculture, home economics and youth work, and have reached mainly the rural and middle-class strata of society. Today, the trend is to fuse agricultural extension with adult education, thereby creating a broader scope of program activities and a means of reaching a wider audience, especially low-income and minority groups. The ultimate aim is to develop community resources and to improve quality of life (Beal et al., 1971).

The ENP is one way through which Iowa Cooperative Extension Service contacts low-income people. The operational objective is to improve nutritional intake and practices of the families through education, thus
improving their level of living. The ENP uses the person-to-person approach through paraprofessional aides, recruited from the neighborhood. The decision to train paraprofessional aides within extension was based on the realization that the characteristics, values, and attitudes of low-income families prevent them from seeking and accepting knowledge and advice from the middle-class oriented extension worker. Researchers generally agree that the paraprofessionals are effective in reaching this kind of audience (Brand, 1971; Nelson & Jacoby, 1973; Rogers, 1973; Synectics Corporation, 1971; Yerka, 1974).

Although the aides from the start of the ENP program were to teach nutrition only, it was soon found that they had to broaden their teaching in order to improve the level of living of their clienteles, i.e., to include sanitation, money management, food buying and storage. Moreover, a nationwide survey (Brand, 1971) revealed that nearly two-thirds of the paraprofessional aides were being asked for family planning or birth control advice by the women and girls in the program. Most program aides saw a need for family planning education among the clients, strongly favored family planning and birth control, and would be willing to teach homemakers about these subjects if they themselves were trained to do so.

Empirical information

A study of available materials and interviews with decision-makers revealed that before the objectives of the pilot project were stated, little empirical information was gathered. As the project unfolded, information was collected as deemed necessary for decision-making. Demo-
graphic data for the area were obtained from the 1970 census and studied
together with routine data gathered in the ENP program.

Informally, observations by the local ENP home economist suggested a
need for family planning education within the ENP program. Such observa-
tions included the high number of one-parent families participating in the
program, and remarks from the aides that they received questions about
family planning and health problems that are usually associated with
having a large number of children. Further support for the pilot project
came from the local family planning clinic itself, where the aides regu-
larly worked in their capacity as ENP aides. The clinic expressed a con-
cern about not being able to reach low-income clients to the extent it
wished to do.

Stated objectives

Context evaluation serves planning decisions to determining objectives
(PDK Committee, 1971). The objectives for the pilot project were identi-
fied (Note 1, pp. 4-5) as follows:

The Family Planning Education Project (FPEP) intends to:

1. Widen the scope of the ENP and through the incorporation of areas of
family life education to include such areas as health, nutrition,
human development, and others as necessary.

   a. Broaden the sources of information of the homemaker with the pur-
   pose of providing a wider range of choices and alternatives in
   the use of family resources.

   b. Make possible more rational decisions in relation to family
   planning.
2. Provide the opportunity for the home economist to use the broad range of her training in such areas as health, nutrition, home management and human development in approaching the topic of Family Planning-Decision-Making in the context of the total family well being.

3. Provide family food aides with the appropriate information.
   a. Develop skills which will allow them to create awareness in the homemakers of their capacity to determine the number of children they want to have and when to have them.
   b. Provide information to the homemakers of the services available in the community for family planning.
   c. Assist the families with other family living related information which could facilitate the decision making process in family planning preference.

For evaluation purposes, all objectives except 3a and 3b are too general and too broadly stated to provide a base for objective and empirical information gathering. Even if more specific objectives were stated, it would be very difficult to attribute cause-effect to the program in question as the aides with their weekly to monthly contact with each client are only one of many sources of influence and information to the families. Some concepts, such as "rational decisions in relation to family planning" would have to be more closely defined as they are intimately tied to cultural and personal values.

Summary

The decision to develop a pilot project in family planning in the ENP program was made at the higher administrative level of Extension Home
Economics and the College of Home Economics. It was based on a need to try out the effectiveness of integrating family decision-making with home economics subject matter and family planning education; to find methods, techniques, and strategies efficient in reaching low-income and minority groups with family planning education; and an observed local need for such a program. Objectives for the program were identified.

Input Evaluation

The purpose of input evaluation is to provide information for determining how to utilize resources to meet program goals (PDK Committee, 1971, p. 222). This is accomplished by identifying and assessing relevant capabilities of the responsible agency, strategies for achieving program goals, and designs for implementing a selected strategy.

The PDK Committee (1971) points out that when the change setting is homeostatic, i.e., where small change is needed and where much information is available to support it, little formal evaluation is usually required. When less information is available or where changes are great, formal input evaluation becomes crucial. In the program under investigation, little information was available about the innovative part of it, whereas organizational changes were small and much known about the ENP program of which the pilot project was a part.

One main question and two detailed ones were raised under the heading of input evaluation, viz.:

What resources were available, and how were they to be utilized in order to meet program goals and objectives?
1. What human and nonhuman resources and obstacles within and outside the organization were identified during the planning stage?
2. What strategies were selected on what basis for achieving the program goals?

**Identified resources and obstacles**

The pilot project consisted of broadening the scope of the ENP program, hence many of both human and nonhuman resources were common to both programs. Most of the information about the innovative part of the program that was used in the decision-making was collected through informal discussions between the assistant state leader and resource persons in extension and at the university.

Resources available were the existing ENP program within the Cooperative Extension Service organization; positive attitude of state and local level administrators; state level extension specialists in home economics subject matter areas as well as specialists in such areas as sociology, health education, education, and evaluation; a local ENP home economist, bilingual, with cross-cultural experience and looking for ways to expand the program. There was a possibility of obtaining funds for the project outside the ENP program. Two communities were available for the project, both with a (for Iowa) high percentage of minority groups and, being urban areas, having a concentration of low-income families. In both communities, family planning programs were established and wishing to reach more people in the low-income category. In general, good working relations among organizations in the communities already existed.
An effort to pinpoint home economics extension's competencies to become involved in family planning was made in October, 1974, when the project was under way. Four state extension home economics staff members identified organizational and subject matter competencies. Identified limitations were the traditional role of the home economist as perceived by staff and community, as well as the lack of awareness on the part of staff and supporting groups of the need for family planning as part of the program, and as a need of low-income people.

The above effort was a necessary step when planning a program. It might have more impact and save time and efforts if executed at an earlier stage before strategies were decided upon, if the scope had been widened, and if more people had been involved. On the other hand, as the nature of the program was developmental and constantly changing, it was at times difficult to separate planning and process stages of the project.

When selecting a strategy, usually identified obstacles can be counteracted and resources maximized. Possible obstacles for family planning education include such things as cultural values and personal attitudes of clients, aides, professionals, and the community-at-large. Values and attitudes take a long time to change.

Much of the communication about family planning is using a negative approach, trying to create fear in the receivers, which in turn would make them abstain from having more children. More people to share limited resources is representative of this approach. Attitude and behavior changes are more likely if fear appeals are not utilized in a message (Rogers,
Emotions in a receiver raised by fear appeals may cause him/her to ignore the message content or the unpleasantness may cause hostility toward the source of the message.

Working against family planning education is also the relative taboo the communication carries and the closed network in which it usually occurs. Among low-income people, family planning education ideas seem to be less taboo than anticipated (Rainwater, 1960), in spite of the fact that the more personal the level of reference to the topic, the more taboo the communication (Rogers, 1973). Taboo communication occurs between very highly homophilous individuals (Rogers, 1973). The aide, being recruited from the neighborhood and viewed as a "significant other" once she has established and proven herself with her client, is in an excellent position in this respect. At the same time, as she works with many people, she could help to break up the closed network of the communication, get it talked more freely about, thus lessening the taboo nature of the topic.

Furthermore, negative rumors and contradictory messages are widely circulated in regard to family planning ideas. They are emotionally laden and difficult to counteract with even factual information. One, for example, is family planning associated with fear of genocide by some ethnic subgroups (Turner & Darity, 1973; Weisbord, 1973). Another example is the firm belief of individuals that contraceptive methods can cause a wide range of sufferings, cancer being one of these.
Selected strategies

Intentions of the program  The intention of the pilot program was motivation through education, approaching the problem of family size in relation to the integrated needs of the family. Thus, once a client became interested in limiting her/his family size, the aide would refer her/him to other agencies for specific help, rather than to provide this herself.

This approach emphasizes the educational nature of the program. It necessitates a close cooperation with other organizations, which in the long run might provide a means for better utilization of community resources.

Site selection  The originally selected site for the pilot project was an urban area with some industries, the main one related to agricultural crops. It had an influx of Spanish-speaking migrant workers. One of the aims of extension is to reach minority groups, and because of differences in culture and language, it was believed that this particular group could benefit from the program. The ENP home economist was bilingual and had experience in working in a Spanish culture. The ENP program was already introduced in the urban area, albeit not firmly established, and a family planning clinic and a health center were in operation.

Staffing  Funds outside the ENP program had been secured for the employment of two full-time aides. Because of integration with the ENP program, it was decided to distribute the funds so that more aides would be employed in the project, spending part of their time in the pilot
program and the remainder in the regular program. From the outset it was decided to involve all existing aides who so desired in the additional training needed for the project. A final selection of aides was to be made when the training was completed.

The local ENP home economist as well as the assistant state leader were to spend one-tenth of their time with the pilot project.

Job descriptions for staff were drafted by the assistant state leader, except for the aides. This was done by the ENP home economist.

Planning committees A state planning committee was established in March, 1974, composed of administrative staff, specialists, consultants, and the local ENP home economist. The committee was appointed by the assistant state leader, who also proposed the functions of the various members. The purpose of the committee was to plan strategies for the project and to facilitate coordination between the different subject matter areas and approaches.

On the local level, a committee was organized, consisting of the local home economists, one area specialist, and the area extension director. The local ENP home economist served as a link between the state and local committees.

Legitimation An important step in initiating a new program is the legitimation of the project by key people in the formal as well as in the informal legitimation structure. Legitimation, or sanction for action, in the planning process consists of consultation with formal and informal leaders of the groups and agencies within the relevant social system (Beal et al., 1966).
Legitimizers do not always commit themselves to help initiate or carry on the new program, but if bypassed, they may throw all of their resources into the blocking of the program. Furthermore, legitimizers constitute a valuable sounding board and often contribute to the clarification of thoughts on the part of the initiators.

Within the formal legitimation structure, the state leader and the local ENP home economist cleared the pathway for the project. During the process stage, the extension specialist in sociology, a member of the planning committee, made a survey of existing state level organizations with family planning or related topics in their program, and thereafter contacted these agencies for mutual information and clarification of purposes. At the local level, the ENP home economist and the assistant state leader, in close cooperation with the extension sociologist, performed the same tasks.

**Audience**    The audience of the pilot project was intended to be recruited from the regular ENP program, from families already involved by the aides. Women, men and youth were to be involved. Based on the aides' performance in the ENP program, it was estimated that each aide could work with approximately 16 persons at any one time.

As the program aimed at an expansion of the ENP program, the aides would use the same approach to the clients, i.e., working on a person-to-person or small group basis, using informal conversations, and utilize the "teachable moment" to introduce the new concepts. Depending on each situation, the aide would have to determine where to start with each individual.
Training of aides. The strategy for training the aides would follow the usual extension pattern, with the state specialists training the local home economists, who in turn trained the aides.

Before the training, an investigation of the training needs would be done through an inventory of the knowledge and attitudes about family planning of the aides. State specialists in a visit to the site were to poll the perceived needs by the local home economists and at the same time through informal talks with the aides tap their needs of training.

In order not to interfere with the regular ENP program, it was planned that the training of the aides be spread out over a period of time, with an average of one training session per month.

Due to the newness of the concept of family planning and its integration with the many areas of home economics subject matter, the state planning committee would serve as a coordinating body between the specialists, when developing teaching strategies and materials.

Visual aids. The nutrition newsletter used in the ENP program had been found to be quite effective in terms of gaining the clients' attention and in generating interpersonal communication, although to have no impact on adoption of new nutrition behavior (Groves, 1973). The aides themselves felt the newsletter to be a good tool in approaching the clients. It was therefore predicted that some visual material might be needed in promoting the new program. Thus each state specialist would develop a short, simple brochure in her area of competence. These would later be used by the aides in their work with the families in order to
explain the program as well as for motivational purposes. Materials from other educational program, such as Planned Parenthood, would be used as well.

**Summary** On mostly an informal basis an appraisal was made of human and nonhuman resources and obstacles to the project. This information, together with the objectives of the pilot project, formed the base for deciding on strategies for reaching program goals. Thus decisions were made in regards to intention of the program, selection of site, staffing, audience recruitment and approach, legitimation of project, training of aides, and development of visual aids.

**Process Evaluation**

Within the framework of the CIPP model, process evaluation has three main objectives. The first is to detect or predict defects in the procedural design or in its implementation. The second is to provide information for programmed decisions, and the last is to maintain a record of the procedure as it occurs (PDK Committee, 1971, p. 229).

Process evaluation is more important than product evaluation in the early stages of program development, but this trend reverses as program development designs become less exploratory and more structured. There is an interdependency between process and product evaluation in the sense that process evaluation is needed to aid in interpreting outcomes, and that need for changes in the present process cannot be properly determined without knowledge of what effects it is producing (PDK Committee, 1971).
The committee further asserts that especially in the incremental and neomobilistic decision settings, a full-time evaluator, instruments for describing the process, regular meetings between the evaluator and project personnel, and frequent updating of the evaluation design are four essential ingredients of process evaluation designs. The present study violated these requirements, which constitutes one of the limitations of the study. Nevertheless, it was anticipated that in fulfilling the conditions as well as circumstances permitted, some insight in the process might be gained that could be of value to the program decision-makers.

The main question asked of Process was: What were the main features of the project design, and which ones contributed to the success or failure of the design to reach program goals and objectives?

This general question was narrowed to six specific questions, viz.:

1. What changes were made during the implementation - based on what information, what alternatives, using what decision rules?

2. What kind of leadership was provided? Expected?

3. Did people involved in the program understand and agree with the intent of the project?

4. What training methods were successful with the paraprofessionals - as viewed by the paraprofessionals, their supervisor, and the trainers?

5. What training was provided for the local home economists? For the state specialists?

6. What approaches did the paraprofessionals use that they consider successful or failures?
Changes during the implementation of the pilot project

Two things should be kept in mind when considering the changes that occurred in the program during its initial stage. Firstly, the very nature of the program was developmental, which in itself denotes constant changes. Secondly, the emphasis of the project during the time data were gathered was on establishing the program and on the training of personnel.

Intention of program  No changes occurred during the study in the intention of the program. The thrust remained motivation for family planning through education. Family planning was interpreted as the decision by individuals to have the number of children they want and when they want. Decision-making thus was a main concept, as was family planning communication. A change occurred in this framework when it became evident that the aides found it difficult to identify the stage in which a family was in the decision-making in regards to family planning. Thus the emphasis became the identification of needs and problems of the participants and then to relate these to concepts in home economics and family planning. The belief remained that motivation for family planning could be achieved if the individual realized the relationship of family goals, family resources and family size.

To emphasize the educational aspect of the program, it was decided at a state planning committee meeting to call the project Family Planning Education Program (FPEP).

Site  From the beginning an urban area in southeast Iowa was selected as the site for the pilot project. Reasons for this choice are described under Input evaluation. In April, 1974, a decision was made to
launched the pilot program simultaneously in a second, neighboring urban area. Due to the unsteady paraprofessional aide situation in the first community, and the lack of Spanish-speaking aides, plus the fact that the ENP program was not very strongly established, made it less likely that a pilot project would succeed. Furthermore, a change in leadership in the family planning clinic had left a temporary absence of a decision-maker in that program.

In the second site, the ENP program was firmly established and had several aides with a comparatively long employment record. One aide had previously worked with a family planning program. The ENP home economist, in charge of the programs in both sites, had personal knowledge of and cooperation with several agencies outside extension. For example, the ENP program already provided nutrition education in the waiting room in the family planning clinic and in the place where the Food Stamp program was administered.

The state planning committee in its second meeting was informed about/decided upon/approved the change. The choice of words in the above sentence depends on the individual committee member. Further discussion of this will occur under the headings of Leadership and Planning committees.

The choice of a site for a pilot project is crucial. Conditions as favorable as possible are sought for the program in order to give the program an opportunity to become established. All social action and change takes place within the context of an existing social system. It is necessary for change agents to have an understanding of the general social
system, the important sub-systems, and the extra-system influences upon the general system and the sub-systems. This will enable them to analyze external and internal directional forces in a situation that will aid or hinder the introduction of a new program. If opposing forces to a specific program are equal to or greater than supporting forces, the program will not be a success. Such an analysis requires a thorough knowledge of the community, which is costly in both manpower, time, and money to acquire; besides, it is rapidly outdated.

From an evaluation point of view, even where conditions are favorable for a program, it is difficult to attribute positive or negative changes to that particular program. The more confounding variables that enter in, the greater the difficulty to establish cause-effect relationships, and consequently, a higher risk-taking in recycling decisions.

Staffing Four aides for each community, or a total of eight, were selected before the project was taken to the field in September, 1974. The aides were to spend one-fourth of their time in the FPEP program and the remainder in the ENP program. The training was not completed at that stage, but the aides were impatient to begin their work. If not allowed to do so, there was a risk of their losing interest in the program.

During the time of the study, there was a constant turnover of aides. At the end, only two original aides remained. The reasons for the turnover varied, but in general did not relate to the nature of the program. At the start of the training, aides could choose to participate in the FPEP program or not, and as new aides were recruited into the program, the nature of the project was made clear to the applicants. Personal reasons,
such as moving from the community or marriage, were often given by the aides. Better paid job opportunities were available, especially in the more industrialized second site.

The rapid and constant turnover of aides created a situation where the need for repetitious initial training arose. Further discussion of this will be found under the heading of Training of aides.

Throughout the study, the assistant state leader and the ENP home economist continued to spend one-tenth of their time in the program. No turnover occurred.

Local home economists had approval from local administration and advisory councils to be involved in the program. In some instances, uncertainty about how much time to spend on the project occurred. Generally, local home economists felt their main responsibilities were to their own programs, but they were all interested in assisting the pilot program. From the outset they perceived their role as resource people to the program, receiving training themselves and information from the state specialists as needed. As the project proceeded, the home economists lost sight of their role in the program and began to question the worth of their involvement. Once this became apparent, regular meetings with staff members as well as involvement in the training of aides became scheduled, which restored the equilibrium.

Planning committees The state planning committee, composed of administrative staff, specialists, consultants and the ENP home economist, met regularly on mostly a monthly basis. Thus nine meetings were held from March, 1974 through June, 1975. The function of the various members
as envisaged by the assistant state leader (Note 1, pp. 11-12) paints a picture of a body of leadership.

Only one of the specialists and local home economists interviewed in this study mentioned the state planning committee as a policy making body. One specialist thought it should make policy decisions, but was glad when this did not occur. Reason given was she felt uncomfortable making such decisions about this project. Examples of responses from the participating home economics specialists to a question about the function of the committee are: "sounding board for ideas", "information" and "decision on action".

An analysis of the agenda for the meetings showed a heavy emphasis on information giving, i.e., framework of project, progress reports of project, reports from specialists, results of questionnaires. Part of the information given at these meetings was with the covert intention of in-service-training, as the members had little experience in working with low-income people and groups with different ethnic background, as well as in the decision-making approach to family planning.

In response to a questionnaire sent by the assistant state leader to the state specialists and local home economics staff, just over half of those responding listed their involvement and participation in the committee meetings as the most frustrating thing in their work in the project. Slightly less than half listed time spent in the committee meetings as most frustrating.

At the state committee meeting held in June, 1975 (the last one included in this study), possible changes in the ways the committee worked
were discussed, but apart from meeting less frequently, no changes were later implemented.

The local planning committee met "when needed" in the beginning. When it became apparent that the home economists required more involvement in the project, regular meetings were held to inform about progress of the project and to discuss related matters.

It appeared that the planning committees served a necessary and useful function of coordinating different subject matter areas and approaches to training, of keeping the members informed about advancements of the project, and as an in-service-training opportunity. On the other hand, they also constituted a source of frustration because of the time involved and of an expectation of greater personal involvement and contribution to the meeting.

**Legitimation**

Legitimation was carried out as planned. Extension councils and Home Economics Education Committees in the counties involved were contacted, as were family planning centers, health centers, a maternal health center, a migrant center, the Social Services and the Visiting Nurse Association. At the state level, more agencies were identified to be involved in family planning education than was anticipated. Thus it was not possible to contact all of these within the comparatively short period of time.

**Audience**

The audience of the project was intended to be recruited from the ENP families and to include women, men, and youth. No change in the program occurred in this respect. In the initial stage mostly women were involved, and to some extent youth, but the men were scarce. The latter might have been due to the fact that in the beginning the aides
were barely comfortable in discussing family planning with women. The taboo nature of the topic still clung, and it took time to get used to using taboo concepts in conversation in a natural manner. Besides, the aides worked during the day when many men are at work and the approach of the aides to men might have to be different, as men often are concerned with other aspects of life than are women. If so, the aides would need to receive training in this area as well.

A question that immediately arose was about criteria for enrolling a family in the FPEP rather than in the regular ENP program. It was decided by the ENP home economist and the assistant state leader to include a family in the FPEP when the aides had introduced the topic, as evidenced by their written logs. No decision was made as to criteria for graduating a family from the program.

**Training of aides**  
Change occurred in the training pattern. Usually specialists train local professionals, who in turn train the aides. It was suggested by the local home economists that state specialists train the aides directly, with the home economists joining the group. This way, both state specialists and home economists would gain experience in working with low-income people through the interaction with the aides. Thus for the first sessions, specialists spent some time with the home economists alone, and then trained aides and home economists together.

This approach was abandoned after both groups voiced their dissatisfaction with the arrangement, the professionals because of the amount of time involved and the paraprofessionals because they did not feel free to express themselves in a mixed group. Furthermore, they expressed a
feeling of being used for the purpose of training the staff, which they
resented. From then on, specialists met with the two groups in separate
sessions. The local home economists became involved by doing follow-up
training of aides, or as in one instance, by doing the training of aides
in the presence of the specialist.

In the beginning there was no follow-up of the initial one-session
training, except from clarification sessions by the ENP home economists in
the regular ENP meetings. Weekly two-hour meetings with the FPEP aides
and local resource persons were started during the fall, and local home
economists, family planning social worker, family planning nurse met with
the aides. These meetings ranged from informal discussion, or "rap
sessions", to planned additional training with emphasis on the first.

The turnover of aides raised the question of ways and means of ini­
tial training of new aides. During the study, new aides were trained by
the ENP home economist until they caught up with the ongoing program.
Taped sessions and visual aids from the original training were utilized.
If the turnover continues, and this is likely according to studies of
paraprofessional aides, it might be worth the time and effort to look into
alternative ways of training joining aides.

Visual aids No changes occurred in regard to plans to produce
pamphlets to be used in the project. A promotion brochure, explaining the
project to the clients, was developed, as was one for motivational pur­
pose and a series of family planning education leaflets dealing with sub­
ject matter areas. State specialists prepared these educational aids
which were circulated to committee members and aides for comments and
alterations before final printing. Late in the study it was found that
the aides were not utilizing the brochures developed in the project. It was not possible at this stage to probe into the question, but it would be of interest to examine the aides' reasons for preferring educational material developed by other agencies. It is possible that the clients are in the knowledge stage in the Innovation-Decision process described by Rogers (1973, p. 103) and not ready for persuasion, the second stage in the above process.

Leadership

Leadership involves accomplishing goals with and through people. Styles of leadership vary, ranging from the authoritarian one in which the leader's authority is inherent in the position he/she holds, to the democratic where authority is granted the leader by his/her followers, to the laissez-faire style where no leader can be distinguished in the group.

No one leadership style is "the best" in all instances. The method or combination of methods that will produce goal attainment depends on such factors as the rank and personality of the leader, the situation or problem at hand, and the type of group and group members involved. A successful leader varies his/her leadership accordingly.

What is important is that there is a congruence between expected and provided leadership style. An incongruence in this respect is apt to cause delay, if not a complete halt, to activities leading to the achievement of stated goals.

Extension is a bureaucratic organization with a well-defined hierarchy of formal leadership. Authority is in the position, being delineated in the job description available for each position.
It is therefore not surprising to find that expected leadership for the project on the part of the state specialists and local home economists was the traditional, formal leadership, based on authority in the position, coming from the top and being disseminated through established channels of communication. Although the state specialists and local home economists felt free to communicate ideas upwards, they would only in exceptional cases bypass an established channel.

The assistant state leader, comparatively new to extension, expected to perform a more democratic leadership role, with its freer flow of ideas and greater volume of communication among the total group. In particular, as the ENP home economist was in a position to follow the development of the project closely and was familiar with the local conditions, the assistant state leader expected more initiative to come from the local level, whereas the ENP home economist was waiting for directives. Furthermore, the assistant state leader had envisaged that the state planning committee would execute a high degree of leadership in the project, but none of the interviewed home economics specialists perceived the committee as doing so.

Informal leadership through opinion leaders is often as effective as formal leadership, and in many instances more so. In this study little effort was made to pinpoint opinion leaders, but it is probable that the ENP home economist, due to her personality, would emerge as an informal leader in several of the groups involved.
Understanding and agreement with program intent

Individuals, organizations, and social systems have a tendency to resist change. Organizational resistance may stem from the individual personalities that make up the organization, the organization itself, and the culture in which the organization operates. One way of lessening the resistance and thus allowing for efforts to be pooled and directed towards a common goal is to make sure people involved in a program understand and agree with the intent of the program. In other words, the different levels of personnel need to adopt the new idea themselves before they can contribute effectively to its further diffusion, and ultimately the program audience needs to understand and agree with the program intent if they are to adopt the innovation.

Communication becomes very important in this process. Communication theory states that in order for persuasion communication to be powerful, the source of the message needs to be perceived as having both competency and safety credibility by the receiver. Channels need to be the ones the receiver usually attends to. The message needs to be coded in a way that it can be decoded by the receiver without loss of intended meaning. Furthermore, communication is more effective between persons who are alike in such attributes as beliefs, values, education, social status, or where a homophilous communication relationship exists. In order to facilitate adoption, the receiver needs to perceive an innovation as having certain characteristics. These include relative advantage over the idea it supersedes; compatibility with existing values, past experience and needs of the receiver; trialability, or possibility to be tried out on a small
scale basis; observability, or visability to others. Complexity, or difficulty to understand and use of innovation acts as a negative force against adoption.

The communication in this pilot project was a multi-stage flow one, i.e., it passed through many levels of personnel with each level functioning as receiver/channel/sender. Program clients would be the last link to receive the message, sending feedback messages. Messages in a multi-stage flow situation are more likely to have a high degree of deletion, distortion, and reinterpretation than in a dyadic one.

Did the persons involved understand and agree with the intent of the project? Yes, and no, and depending on whom.

In the beginning state specialists immediately agreed with the intent but had some difficulties in understanding how their subject matter area could relate to family planning. As their contact with and experience of low-income families was rather limited, they were hesitant in how to approach this audience, even though done through the aides. As the program developed and changed, the understanding of the intent decreased. The monthly planning committee meetings served as information sources of what happened, but did not aid in understanding the intent. At least one specialist expressed a reluctance to ask for clarification in an ongoing meeting. In the beginning, written material about decision-making, population, and cultures was distributed and/or suggested for reading, but due to heavy work loads, the specialists could devote only a limited amount of time to the project.
As far as the state specialists were concerned, the source credibility was evident, a heterophilous relationship existed between source and receivers. The innovation was perceived to have relative advantage and compatibility by the specialist but observability was lacking, and the complexity great.

There was a breakdown in communication which hindered at least temporarily the understanding of the intent of the program on the part of the state specialists. Lack of time of both sender and receiver was a major contributing factor to this.

The local home economists agreed with the program intent when exposed to the innovative idea. Understanding was not complete in the beginning, and like the state specialists they had difficulties in seeing how family planning, decision-making, and home economics subject matter areas could be integrated. This was not a source of frustration because of promised training and involvement in the program, which would ultimately bring about understanding. But as the program progressed, and changes took place, frustration occurred. The local planning committee did not meet regularly, due to difficulties in finding a common meeting time. The involvement of the local home economists in the program became less than expected due to changes in the training plans. The frustration level became high and after a confrontation, regular meetings of the committee were scheduled to keep staff informed about the project.

From the beginning of the project, the aides in general neither overtly agreed nor disagreed with the program intent. One sign of disagreement might have been the reluctance on the part of black aides to
voluntarily participate in the training and refusal to participate in the program. The reason for this was not established. One aide, who had previously worked in a family planning project, was in favor of the program. Source for the initial message was the ENP home economist, who had source credibility. Later on, when training was led by state specialists, and with the local home economists present, the intent of the program became less well understood. The aides voiced their concern to the ENP home economist. As source of message, the state specialists had competence credibility, but lacked safety credibility. Specialists and aides were heterophilous in regards to their communication relationship, as were the ENP home economist and the aides, but in the latter case, empathy lessened the effect of the heterophilous relationship.

As time went on the training progressed, and as aides started to work within the project, the understanding and agreement with the project intent increased. One contributing factor might have been that although a turnover of aides occurred, there were always a few who could serve as a homophilous source of information about the project intent.

Yet on the whole, there was a communication breakdown about program intent. In the beginning of the project, all involved with the possible exception of the aides, overtly agreed about the program intent. This agreement might have been based on lack of full comprehension of the intent (but with the hope of fuller understanding as the project developed), paired with felt organizational expectation to participate in the program. Understanding of the program intent was low in regard to how, rather than to what in the beginning. The latter deteriorated after a while. Contributing to this were factors like time shortage, heterophilous communi-
cation relationships, changes in program without additional clarifying or informational messages, and over-use of one channel (like written material to very busy people).

Training methods

A variety of training methods and techniques were used throughout the training, depending on the personality of the trainer and the nature of the topic. Thus no direct comparison could be made between methods. Nevertheless, by observing the reactions of the aides, from remarks made by aides about the training to the ENP home economists with whom they speak freely, as well as from responses of the aides to a questionnaire used in evaluating the first year of the project, some inferences could be made.

The integration of decision-making, family planning, and home economics subject matter was a very complex innovative idea, and constituted a new way of thinking on the part of all involved. This in itself suggested a slow rate of adoption.

The multi-stage flow of communication about the new idea increased the deletion, distortion, and reinterpretation of the message itself, due to selective exposure, selective perception and selective recall at each stage on the part of the receivers. To combat these disruptive forces, repetition of message and more information aimed at increased comprehension, but above all changes in attitudes and beliefs, can be effective tools to use.

The aides, being adult learners, came to the training sessions with past experience and already formed beliefs and attitudes. Therefore, it
was unrealistic to have expected them to comprehend and adopt the new ideas in one half-day training session. Redundant messages over a period of time were needed.

When trainers have just one opportunity to meet with a group, it is human to try to transmit as many ideas as possible. In doing so, the trainer runs a risk of producing information overload, which occurs as a result of excess information inputs over what the receiver is able to process, and utilize. One human expression of information overload is poor performance.

Communication is often defined as a transfer of meaning. Messages can be transmitted, but meaning is in the individual, brought about by internal responses to a stimulus. People will always respond to a stimulus in light of their own experiences. When people are heterophilous, the selection of stimuli, or words, becomes very important if transfer of meaning is to occur. A message can be widely distorted, especially while passing through a multi-stage flow. It is possible that if the language used in the training sessions consisted of words and concepts the aides used in their daily life and work, the comprehension and therefore the rate of adoption would have increased.

The adult paraprofessional aide is less interested in theories than in information that will help her solve present problems. Low-income people live with constant crisis situations and tend to be present, rather than future oriented. Thus training, geared towards the daily living of the aides and clients would tend to be perceived by the aides as relevant. Also, the aides were eager to try out their newly acquired training as soon as possible and not to wait until the training was completed.
The aides brought a wide range of background experience, which they were willing to share. Informal discussions seemed to animate the aides, although there was a tendency on the part of one aide with a strong personality to dominate the other aides. Long sessions forcing the aides to sit still, and the use of sophisticated visual aids tended to be ineffective techniques to use with the aides, whereas role playing, or the viewing of a film followed by discussion seemed to stimulate their thinking.

**Training of state specialists and local home economists**

The state specialists and local home economists had a limited knowledge of and experience in working with low-income people. When the ENP program started in 1968, state specialists received training in doing so, but only one of the specialists had participated in this. Both state and local staff had occasionally been involved in the training of the ENP aides.

As the project used a framework of decision-making and communication principles in family planning, the specialists and local staff needed training in these areas. The integration of family planning with home economics concepts constituted another need for training.

The state planning committee meetings were used by the state leader to give training to the state specialists in the first two respects. Techniques used were illustrated presentation, printed material, films, videotape from training sessions with aides, and feedback about the reactions of aides to training received. Interaction with aides and an occasional visit to a home gave state specialists some insight in working with low-income people.
Local home economists were exposed to the framework of the project at a special training session and through participation as audience in the training meeting for the aides. Through the latter they were also exposed to cultural awareness training and health/family planning training. Thereafter, state specialists held specific training for local staff in subject matter, coordinated with the training of the aides.

Thus both state specialists and local home economists received some training in the project. As the framework of the project was a complex one, it is possible that further training in this respect would have lessened the uncertainty level of the staff and reduced the level of frustration both groups experienced.

Both groups voiced a concern about their lack of understanding of the culture of low-income people and minority groups. As empathy with an audience lessens the effects of heterophily, it might be worthwhile to provide continuous, long-range training in these aspects. Yet involvement in the FPEP program constitutes only a fraction of the daily workload of state and local staff, which limits the time that can be set aside for such training at any one time.

Approaches used by paraprofessionals

Once the aides had gained confidence enough to be at ease when talking about family planning, they did not find it difficult to introduce the topic. Sometimes the homemaker would bring up a question, or the situation would naturally invite remarks that could lead to family planning education. Counseling, or helping in problem solving, rather than telling the clients, seemed to be an approach the aides found successful.
The closed network in which family planning communication occurs was evident in the internal evaluation of the project. The aides said they had communicated their training to their daughters, friends, and neighbors. They overwhelmingly found it relatively easy to talk about family planning education to homemakers they had been working with for a period of time and established a confidential relationship with.

Families seeking information or those feeling they could not afford more children were also comparatively easy to approach.

Summary In the process evaluation changes during the implementation of the pilot project in regards to the strategies decided upon as a result of context and input evaluation were examined. These included the intention of the program, staffing, planning committees, legitimation of the project, project audience, training of aides, and visual aids developed in the project. An exploration of leadership provided versus expected was made. The understanding and agreement with program intent of people involved was investigated. Training methods used with the aides were discussed, as was the training for state specialists and local home economists. Finally, the approach used by the paraprofessionals to their clients was determined.

Product Evaluation

The purpose of product evaluation is to measure and interpret attainments, not only at the end of a project cycle but as often as necessary during the project term (PDK Committee, 1971, p. 232).

The method of product evaluation recommended by the PDK Committee (1971) includes the devising of operational definitions of objectives and
of measuring criteria associated with these objectives, the comparison of these measurements with predetermined absolute or relative standards, and the rational interpretation of the outcomes, using context, input and process information.

Criteria may be either instrumental or consequential (PDK Committee, 1971). Instrumental criteria are related to accomplishments that are at an intermediate level, and that contribute to the achievement of the ultimate objectives. Consequential criteria are concerned with fundamental conditions being sought. In order to collect the type of information that will be useful, the evaluator needs to find out the criteria a decision-maker will apply when making decisions.

The major question raised in this study under the heading of product evaluation was:

To what extent were the program goals and objectives attained?

Two more specific questions were considered to be of special interest under the main question, viz.:

1. What were the accomplishments of the pilot project as related to the program's short-term and long-term goals?
2. What kind of information was needed for recycling decisions?

At the end of the first year after the official start of the project, the assistant state leader conducted an evaluation of the project. Three main aspects were included: the participation of staff and specialists, participation of aides, effects of the project on the program participants. Information was collected through questionnaires prepared by the assistant state leader, and administered to staff and aides; through the
logs written by aides after each home visit; through informal questioning of the aides, and from statistical information routinely collected in the program.

Throughout the project, forms used by aides and administrators in the ENP project had been modified and revised in order to collect information needed for the development of the FPEP project. Thus logs contained information given to the family, materials used in the visit, educational needs as seen by the aides in regards to family planning, food and nutrition, health, child development, money management, and clothing. Problems mentioned by the family were also recorded. Demographic data were collected by the aide when the family enrolled in the program, and thereafter every sixth month. At the same time, the aide recorded such information as visits to family planning clinics and use of birth control methods. She then estimated where the participant was in the decision-making process in regards to family planning. The latter was found to be too difficult for the aides to identify, which became a contributing factor to the shifting of the emphasis in the project to the identification of needs and problems of the family and then relating these to concepts in home economics and family planning.

Between September, 1974 and May, 1975 the aides made a total of 151 visits to a total of 74 families. Aides leaving the program were the cause behind 12 of the 19 families discontinuing the program. Of the 41 ENP families enrolled in the program in May, 18 were between 24-35 years old and 12 between 17-23. None was over 55, and one under 16. No mother had over six children under the age of 19. Statistics in these respects were not available for the 21 non-ENP families enrolled in the program.
Accomplishments of the pilot project as related to objectives

The objectives of the Family Planning Education Program (FPEP) (Note 1, pp. 4-5) were to:

1. Widen the scope of the ENP and through the incorporation of areas of family life education to include such areas as health, nutrition, human development, and others as necessary.
   a. Broaden the sources of information of the homemaker with the purpose of providing a wider range of choices and alternatives in the use of family resources.
   b. Make possible more rational decisions in relation to family planning.

2. Provide the opportunity for the home economist to use the broad range of her training in such areas as health, nutrition, home management, and human development in approaching the topic of Family Planning - Decision - Making in the context of the total family well-being.

3. Provide family food aides with the appropriate information.
   a. Develop skills which will allow them to create awareness in the homemakers of their capacity to determine the number of children they want to have and when to have them.
   b. Provide information to the homemakers of the services available in the community for family planning.
   c. Assist the families with other family living related information which could facilitate the decision-making process in family planning preference.
From an evaluation point of view, the objectives in their present state do not lend themselves easily to evaluation. For assessment purposes, goals need to be clear, specific, and measurable. Clearly stated objectives facilitate the identification of needed information. Specific objectives make the translation of goals into operational terms possible as well as making them visible. As long as goals are clear and specific, usually a way can be found to measure them.

It is important that developmental pilot projects have built-in flexibility, which broadly stated objectives give. On the other hand, too much flexibility leaves room for endless interpretation, and hence, evaluation becomes an onerous task. One way of handling broad objectives is to state some more specific ones after a thorough study of the project, with the calculated risk of slightly shifting the goals. Another way is to make the study exploratory and open-ended, or a goal-free evaluation. Data-gathering becomes a problem here, as it is extremely difficult to observe and collect data about everything.

This study has not concentrated on product evaluation because the emphasis of the program during the data-gathering was on establishing the program at state and local levels, and on training the aides. Aides started working in the field in September, 1974, although their training was not completed until March, 1975. The turnover of aides also caused a dropping-out of some families from the program. In the beginning, the aides spent some time "feeling their way" with the clients and in getting accustomed to talking about family planning concepts with ease. Hence the pace was slow during the first few months.
Objective number one Did the pilot project widen the scope of the ENP program? Officially, the answer is yes, it did. On the other hand, before the program started, the aides were already being asked by clients about family planning, but it can only be guessed if they gave any answer, and if given, it was not necessarily one that was either factual or related to other aspects of family well-being.

After the program started in the field, the aides were in a position to introduce the topic themselves and to follow it up on a long-term basis, instead of answering one specific, isolated question. Much ignorance and many misconceptions on part of the aides were removed during their training.

In the beginning, the aides tended to concentrate more on family planning itself and in giving information about community resources. As they slowly started to realize the relationship of family size, family resources, and concepts in home economics, they began to carry this over to the families, as evidenced in their logs. This trend was just starting to show up at the end of this study.

Did the program broaden the sources of information of the homemaker, thereby giving her a wider range of choices and alternatives in the use of family resources? Did the program make more rational decisions possible in relation to family planning? No information was available that could provide answers to these questions. No benchmark study was made in this regard. The ambiguity of the objectives and the value-laden concepts involved make such an evaluation extremely hard.
Objective number two  The program did give the home economist an opportunity to use the broad range of her training in approaching the topic of family planning-decision-making in the context of the total family well-being. The opportunity was there as soon as the program started. Did she take it? Assuming that "the home economist" means all persons involved in the program with a home economics professional training, the answer is yes, to some extent.

Because of the new way of thinking that this approach constitutes, all home economists in the beginning were groping around, seeking ways of doing so. Once state specialists found key concepts in their area to integrate with decision-making and family planning, the planning committee meetings provided a meeting ground for the different subject matter areas. Written material about the project that was made available to committee members reinforced this. The same opportunities were open for the local home economists, who had the advantages of getting information from the specialists and of being closer to the local situation. The ENP home economist and the assistant state leader who were closely involved with the project and attended all training meetings had probably been the ones who have come the furthest in integrating family planning, decision-making, and home economics.

Objective number three  The project intended to provide family food aides with the "appropriate information". The term appropriate needs further clarification, and moreover, it was an early stage in the program for formal assessment of the relevance of the information the aides had received. The logs provided data on information given to program partici-
pants in each visit. This could be compared with the information the aides had received in their training. Following a recommendation from the state planning committee, the aides were exposed to general information about population problems, decision-making, and poverty in order to create awareness of the problem at hand and at the local situation. An effort was made to create an awareness of cultural differences in family life-styles and its implication for the program and the aides. Furthermore, specific training sessions were given in health, food and nutrition, housing, money management, child development, and clothing and in their relation to family planning and family resources. A face relationship between information received and given certainly exists, but it might be worthwhile to probe into the effects of incongruence in levels of information received and given.

Another intention of the program was to develop skills on the part of the aides which would allow them to create awareness in the homemakers of their capacity to determine the number of children they want to have and when to have them. The aides received training in counseling techniques. They seemed to find that the approach of helping clients to solve their own problems gave better results than telling people what to do. Through counseling techniques, alternative ways to solve a problem are explored, which would provide a wider range of choices and alternatives in the use of family resources. No formal attempt was made to assess if homemakers were aware of their capacity to decide on the number of children to have and when to have them, but the very fact that clients ask for information about family planning methods are evidence of awareness.
Do the aides provide information to the homemakers of family planning services available in the community? Yes, they do, as evidenced by their logs. It is probably one of the areas where they are most effective because of their knowledge of and understanding of these facilities, gained through close cooperation and experience with the agencies in question. The latter is a side-effect of the program.

Do the aides assist the families with other family living related information which could facilitate the decision-making process in family planning preference? It was found at an early stage that it was impossible to separate out the FPEP part of the aide's work from her ENP part. Therefore she provided information that could be grouped under the headings of: family planning and related topics (VD, drugs), community resources, health-related problems, food and nutrition, child development (problems with adolescents), housing, money management, and clothing. Thus the aides did provide information that could contribute to the above mentioned process. If it will in fact contribute depends to a very great deal on if the aides themselves see the relationship between the concepts involved. Indications at the end of the data-gathering period were, as indicated in the logs and from remarks from the aides, that the aides were slowly starting to realize the connection.

Information needed for recycling decisions

Recycling decisions are used in determining the relationship of attainment to objectives, and whether to continue, terminate, evolve, or drastically modify the activity. As recycling decisions basically involves product control choices, they are concerned with attainments at any
point in the program. Authority for recycling decisions usually resides with the operations manager during the implementation of an activity cycle, and with the responsible fiscal agent at the conclusion of an activity. The policy administrator is a key figure in recycling decisions (PDK Committee, 1971, pp. 83-84).

Decisions are not made in a vacuum, but always in relation to some explicit or implicit value structure. Values and criteria are closely related. An evaluator needs to know what criteria will be used in recycling decisions in order to collect useful information. In this study, no attempt was made to deal with values as it is a very complex topic. Personal as well as organizational values have to be considered, and these can be overt or covert. It takes a fairly lengthy period of time of close interaction before even the overt values of a person become identifiable.

Recycling decisions are concerned with attainments at any point in a program, not just with outcomes following a full circle of a program. At the implementation stage they focus on the extent to which desired ends are being met as opposed to implementing decisions that focus on the extent to which means are operant as intended.

In this study, an effort was made to delineate information needed for recycling decisions in the pilot project. This was done by an inquiry into what information the assistant state leader, in her capacity of program manager, and the state leader, as policy-maker, felt they needed for their levels of decision-making.
The assistant project leader conducted an evaluation after the first year of the project (Note 2). Thus, information that was considered needed for recycling decisions at this stage included the involvement of the staff and specialists in planning committees and training meetings in the field, as well as their suggestions for improvement of their participation. Data on the aides and their work with the clients, on their application of information given in training sessions were needed, and the suggestions of the aides for improvement of their training were solicited. The third area of interest was the effect of the program on the audience, including changes of attitudes and behavior of the participants toward family planning; alternatives and information about family planning given by aides; and action taken by participants toward family planning. Finally, needs and problems identified by aides and their relationship with subject matter areas were to be considered.

The state leader in an interview expressed a need for information that could substantiate whether the program participants were better off after the program than they were before, when making recycling decisions at the policy-making level. Answers to the following questions would be used as criteria for recycling decisions: Are people making rational decisions? Are they aware that they have a choice, and do they have the ability to make decisions in regards to the number of children and the spacing of them? Can they relate family planning practices to the well-being of the family? The number of participants in the project and their use of community resources and facilities, family planning clinics in particular, would also be applied as criteria.
The study of a single program can show whether participants are better off after the program than they were before if a study of variables involved is made before and after a program goes into operation. Cause-effect relationship, or whether the program did bring about the benefits rather than extraneous outside conditions, can be established if a control group or a comparison groups is used. Difficulties inherent in establishing control groups were discussed earlier in the literature review.

Reliable and valid answers to the stated questions would be extremely difficult to find at the present time. Very sensitive instruments will be needed to measure the concepts involved. The difficulties are increased by the fact that the characteristics of the audience makes it necessary to apply unobtrusive and nontargeting evaluation techniques. Future research might solve the problem of merging these two circumstances.

Summary In this study the product evaluation was not emphasized due to the early stage of the program. As process and product evaluation are closely related and give meaning to each other, some effort was made to assess the product. Information was gathered from three sources. These included data collected in an internal endeavor to assess the pilot project one year after its official start, the aides' logs, and routine data in the project. Accomplishments of the project as related to objectives were dealt with, and an attempt to delineate information needed for recycling decisions was made.
It is assumed that program evaluation ultimately will lead to improvements of present or future programs. Through an evaluation study, it is possible to analyze a program in terms of its objectives, the assumptions underlying these objectives, the specific program activities designed to achieve these objectives, the rationale for believing that these activities are capable of attaining the objective, the separation of the "idea" of the program from how it is being carried out, and the determination of criteria for observing the extent to which the objectives are being attained.

Traditionally, a program is evaluated by its product, using the objectives of that program as criteria. In education it often takes a long time to ascertain if the product is acceptable or not. In addition, this kind of evaluation rarely indicates where changes are needed within the program. Because of the rapidly shifting social, economic, and intellectual needs of society, adjustment to changing situations need to be made early in an educational program so as to maintain desired outcomes. Therefore, it becomes important to identify key decision points in a program where the consequences of a decision or changes introduced, can be more easily and rapidly predicted.

An educational program planner is thus faced with the constant task of decision-making. Because any kind of decision involves a choice situation, the decision maker makes a conscious or unconscious evaluation
of available options. Ideally, decisions should be based on sound information about alternative plans and procedures, and not on intuition or limited knowledge.

Today, when the trend is towards bigger educational programs, often federally and state funded, and aimed at social action with the ultimate goal of raising the living standard of many people, policy-makers increasingly feel a need to rely less on subjective and political reasons for making decisions about programs. The availability of objective information makes it possible to make more rational decisions. Therefore, information for decision-making becomes an important factor in program planning and because of the integration of program planning and evaluation, it has also become the thrust of evaluation in the last decade.

The program under investigation was a pilot project within the Expanded Nutrition Program of the Iowa Cooperative Extension Service, which received special funding in order to broaden the teaching scope of the paraprofessional aides to include family decision-making as it relates to family planning, and to integrate this with concepts in home economics subject matter areas: nutrition, health, human development, housing, textiles and clothing. Because it was deemed important to evaluate the introduction of the new concept in order to facilitate future decisions regarding the continuation, expansion and/or termination of the pilot project, the Department of Home Economics Education in collaboration with extension undertook to assess the program. The present investigator served as an internal evaluator for the initial part of the project or from March, 1974 through June, 1975. At this point, the pilot
project had gone through one cycle of development. It had been initiated, a state planning committee was operating, local extension home economists were involved, and aides had been trained to relate family planning concepts to decision-making as well as to the different areas of home economics. They were actually working in the field, and using material developed for this purpose.

The objectives of the study were to record and describe the planning and implementation of a pilot program; to analyze the processes and procedures used in the program; and to make recommendations about processes and procedures, based on the analysis, for other similar programs.

Assumptions underlying the study were that subjects involved in the program planning and execution were willing and able to provide needed information; that the Context, Input, Process and Product (CIPP) evaluation model could be entered at any stage; and that information gained from a pilot project can be useful in other similar programs.

Several limitations beset the study. It was confined to one program only. Because of the restricted time the investigator could participate in the pilot program, major emphasis was on the process evaluation, which in turn was based on periodic written and oral communications instead of day-by-day observations. Neither program participants nor the paraprofessional aides were studied in any depth due to the relatively short time-span of the study and to the initial stage of the project. The number of subjects in the study was small, hence statistical analysis was limited and inferences made outside the specific program were very tentative.
An extensive search of literature was made in order to find a theoretical framework for the evaluation of a pilot program. Furthermore, as the project was limited to one program only and involving few people, there was a need to document findings from the literature.

The theoretical framework for the study was a modification of the Context, Input, Process and Product (CIPP) evaluation model, which was thought to provide a holistic and directive, broad base for evaluation. The modification consisted of including such features as: collecting evidence as broadly and extensively as possible from the goal-free evaluation model and the evaluator functioning as a sounding board for the program manager, giving quick, personal judgments about the program from the responsive evaluation model.

Participants in the study were the extension state leader of the home economics program; the assistant state leader directly in charge of the Expanded Nutrition Program (ENP) and of the Family Planning Education Program (FPEP); six state extension subject matter specialists; five local extension home economists; and the paraprofessionals in the pilot project, the number of which fluctuated around eight.

In order to structure the study, questions that might be answered by the research were formulated by the investigator after a search of the literature and informal discussions with extension personnel. The questions were grouped under the headings of Context, Input, Process and Product. They were submitted to a total of five qualified persons in order to select ten which were considered most important for future decisions about the program.
Data were collected from March, 1974 through June, 1975 using such methods as unstructured observation of state planning committee meetings, informal discussions with the assistant state leader, formal and informal interviews with six state subject matter specialists and five local home economists, as well as through unstructured observations of training meetings of local home economists and paraprofessional aides. Comments from the aides about their training, results from internal efforts to evaluate aspects of the project, and information about the project available in files at the state and local levels were utilized.

The findings were derived at by synthesizing the data collected and were presented as responses to the final twelve questions about the subject. A summary of the findings was made in form of conclusions. These, together with recommendations, can be found in the section that follows.

Conclusions and Recommendations

Conclusions in this study have been reached at through the findings in a single program in its early stage of program development, and with a small number of people involved. The generalizability of the findings are therefore limited to this program only. Nevertheless, some of the conclusions and recommendations are more general in their statement, indicating that support for them has been found in the literature, thereby increasing the likelihood that they may hold true for other, similar programs as well.

Coming through the trauma of birth, the program emerged, survived, and showed promises of reaching maturity. It is therefore recommended
that it be given tender loving care, nurturance, and guidance during infancy and the formative years of childhood so that it may come of age in a fully developed, healthy state, thus being able to make its contribution to society in its full capacity.

The stated objectives for the program under investigation were expressed in such universal terms as to be ambiguous. Once a program has established itself, goal evaluation becomes important as evidence of program achievement is vital for administrators when allocating or seeking funds for projects. It is therefore recommended that clear, specific, and measurable objectives be stated for the project that at the same time allows for flexibility in the program.

It is further recommended that a benchmark study be made of the existing conditions before the project expands. Variables to be studied would be selected in relation to the clear, specific and measurable objectives of the project.

As there was ample evidence of a breakdown of communication within the different levels of the project, especially vertically, it is recommended that the flow and volume of communication in a pilot project within the organization be studied.

Because many daily decisions need to be made in a flexible pilot program it is recommended that formal authority for such decisions be given to the local level, or alternatively, more rapid channels of communication be established.

The state specialists play a key role in the integration of family decision with family planning and their subject matter area. They select
key concepts and establish the relationship. Furthermore, they train both local home economists and aides to see this relationship and to help others become aware of it. In order to be effective in cross-cultural training, a trainer needs in addition to knowledge of subject matter and educational principles to have a thorough knowledge about and insight into own and others' cultures. It is therefore recommended that the state specialists receive training in cultural awareness, in cultural aspects of low-income and ethnic groups residing in Iowa, and in working with other cultures.

Criteria for enrolling a participant in the FPEP rather than in the ENP program need, for administrative purposes, to be more refined in order to distinguish between the two programs. For the same reason, criteria for graduation from the program are needed. It is therefore recommended that these two types of criteria be established.

At the end of this study, women and youth only were participating in the program. It is therefore recommended that ways and means of reaching the men be explored. Furthermore, the program has this far reached early adopters, or persons only needing little incentive to come to a decision about family planning. It is therefore recommended that the program start to investigate strategies to work with harder-to-reach audiences.

The turnover of aides was high during the study. This created problems in training the new aides who needed training in both the ENP and the FPEP programs. It is therefore recommended that alternative ways of training aides be explored.
Because of the complexity involved in the integration of family decision-making, family planning and home economics subject matter areas, combined with the characteristics of the aides as adult learners, it is recommended that in the initial training few ideas be thoroughly introduced at the aides level, and that periodic follow-up sessions be arranged in which clarification and adding of new ideas can take place.

The planning committees served a necessary and useful function of coordinating different subject matter areas and approaches to training, of keeping the members informed about advancements of the project, and as an in-service-training opportunity. They also constituted a source of frustration because of the amount of time involved and of an expectation of greater personal involvement and contribution to the meeting. It is therefore recommended that the meetings continue to serve their beneficial functions but that the members be involved in problem-solving to a greater extent as opposed to choosing between solutions. As they have gained experience with the project they are ready to contribute where perhaps before they were afraid of doing so in ignorance.

Whether applying goal-free or broad-aimed evaluation techniques, both methods require a full-time evaluator who functions as a shadow of the program manager. Scriven (1972) even recommends the use of two external evaluators, working independently until the end of the evaluation. He implies the presence of an internal evaluator as well. Few programs would be willing to carry the cost involved. It is recommended that the evaluator be involved from the very beginning of the project. This would facilitate a better understanding of decision rules used in the decision-
making, and make it possible to delineate information needed, and - perhaps - to collect the information and present it when required.

The identification of obstacles within and outside an organization to a program becomes important when planning implementation strategies as an awareness of them makes it possible to counteract or circumvent their effects. It might have been possible to reduce the amount of frustration that almost all groups experienced if attention had been directed towards the need for greater communication flow in a pilot project. It is therefore recommended that the identification of obstacles, along with resources, to the program be identified at an early stage both inside and outside the organization.

As it was discovered that the aides did not use to any great extent the visual aids that were developed in the program, it is recommended that a formal assessment of their format, content, and approaches be made.

The present study leaves an almost unlimited number of questions unanswered. Implications for further research are infinite. This has been an evaluation study and explorative in nature. Evaluation research becomes the next logical step with its potential of establishing cause-effect relationships. In this type of research there is also a need for reliable and valid data collecting devices, especially where the variables have to be measured in an unobtrusive and nonthreatening manner. Possible side-effects of the program on clients, society, staff and organization need to be identified and investigated.


Cooperative Extension Service, United States Department of Agriculture. Program Performance 1971 Expanded Food and Nutrition Program. (ERIC ED 076 870)


Evans, John W. Evaluating education programs: Are we getting anywhere? Educational Researcher, 1974, 3(8), 7-12.


Kirst, Michael W. The rise and fall of PPBS is California. Phi Delta Kappan, 1975, 56(8), 535-538.


Kohl, John W., & Achilles, Charles M. A basic planning and evaluation model for cooperation in providing regional education services. University Park, Pennsylvania: Pennsylvania State University, 1970. (ERIC ED 053 827)


Roberto, Eduardo L., & Gehrz, Susan L. Evaluating family planning program effectiveness and efficiency. Studies in Family Planning, 1975, 6(2), 37-44.


Rountree, Mary Constance. Effectiveness of the Expanded Food and Nutrition Education Program in producing improvements in diets of homemakers. (Doctoral dissertation, Ohio State University, Columbus, 1973). Dissertation Abstracts International, 34, 5548B.


Stuffelbeam, Daniel L. Toward a science of educational evaluation. Educational Technology, 1968, 8(14), 5-12.


Stuffelbeam, Daniel L. Should or can evaluation be goal-free? Journal of Educational Evaluation, 1972, 3(4), 4-5.


Terhune, Kenneth W. Determining the perceived rewards and costs of family size. Paper presented at the annual meeting of the Population Association of America, Toronto, Ontario, April, 1972. (ERIC ED 097 986)


Voht, Donald E. Problems in evaluating community development. Journal of Community Development Society, 1975, 6(1), 147-162.


Weiss, Carol H. Evaluating action programs: Readings in social action and education. Boston: Allyn and Bacon, 1972. (b)


Wick, John W. Educational measurement: Where are we going and how will we know when we get there. Columbus, Ohio: Charles E. Merrill Publishing Co., 1973.


REFERENCE NOTES


ACKNOWLEDGMENTS

Without the support and encouragement of several persons, this dissertation would not have been produced.

Sincere appreciation goes to Dr. Ruth Hughes, whose patient understanding and guidance prodded the investigator along on the thorny road of research.

Genuine gratitude is extended to Mrs. Maria de Colon for her personal friendship as well as for her encouragement and total cooperation during the study.

Acknowledgments are given to the Iowa Cooperative Extension Service for making this study possible and for contributing funds for travel. Grateful thanks are given to its administration and personnel at state and local levels for their cheerful and unquestioning collaboration. Although no names are mentioned, nobody is forgotten.

At last, but not at least, thanks are due to the members of the writer's study committee: Mrs. Julia Anderson, Dr. Irene Beavers, Dr. Mary Heltsley, and Dr. Robert Strahan, for their interest, counsel, and forbearance during the years it took the writer to fulfill the requirements for her degree.
APPENDIX

Questions that Might be Answered Through the Proposed Study

Context

In what context were the goals and objectives of the pilot program formulated?

1. What theoretical information was collected before decision on goals and broad objectives of the pilot program were formulated? Can key points be identified?

2. What information (evidence of needs, study of community, existing services, etc.) was collected before decisions on goals and broad objectives of the pilot program were formulated? Can key points be identified?

3. What values of the Cooperative Extension Service and of persons involved were identified before formulation of goals and objectives? By whom? How? Which were the key ones?

4. What alternatives were perceived by the decision maker to be available?

5. What decision rules were established? On what grounds?

6. What are the short-term and long-term outcomes of the program?

Input

What resources were available, and how were they to be utilized in order to meet program goals and objectives?

1. What human/nonhuman resources and obstacles within the organization were identified during the planning stage? Which were considered to be key ones?
2. What human and nonhuman resources and obstacles outside the organization were identified? Which were considered to be key ones?

3. What strategies were considered for achieving the program goals?

4. What decision rules were used?

5. What strategy was decided upon?

Process

What were the main features of the project design, and which ones contributed to the success/failure of the design to reach program goals and objectives?

1. What happened in the pilot project?

2. What changes were made during the implementation - based on what information, what alternatives, using what decision rules?

3. Is established hierarchy of authority and procedures in an organization a strength or a weakness for a pilot program?

4. Who are the perceived decision makers in the project?

5. Are there one or more key person(s) for the success of the pilot project? Can they be located? What makes them key person(s)?

6. What kind of leadership was provided? Expected?

7. Did people involved in the program understand and agree with the intent of the project?

8. How does the resource person's perception of their role in the program influence their participation in the project?

9. What kind of personal relationships were established within the project? To what effect?
10. What communication channels were utilized at various levels? To what effect?

11. Are established channels of communication within the organization functionable when a pilot program is being established?

12. What training methods are successful with the paraprofessionals -- as reviewed by the paraprofessionals, their supervisor and the trainers?

13. What approaches do the paraprofessionals use that they consider successful? failures?

Product

To what extent are the program goals and objectives being attained?

1. What are the accomplishments of the pilot project as related to program short-term and long-term goals?

2. What kind of information is needed for recycling decisions?