1981

A Methodological Treatment of the Programatic Use of Real Time Attitude Assessment Techniques

S. R. McNeal
Loyola University Chicago

A METHODOLOGICAL TREATMENT OF THE PROGRAMATIC USE OF REAL TIME ATTITUDE ASSESSMENT TECHNIQUES

by

Sherry R. McNeal

A Thesis Submitted to the Faculty of the Graduate School of Loyola University of Chicago in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

August

1981

ACKNOWLEDGEMENTS

Special appreciation and acknowledgement are given to the faculty advisors serving as committee members of this thesis. They both have been so very thoughtful, concerned, supportive and generous with their time and help.

In addition, the special friend who has shared the development of this thesis, in both preliminary and final stages is acknowledged for her generosity in all elements of this project's creation.

FOREWORD

In the summer of 1978, having completed my first year in graduate school, I went out to California to study an exciting field - space colonization. A six week course was being offered in the Sociology Department at California State University, Northridge, and its agenda suggested strong potential for a new area of research in Applied Social Psychology. During the course of that six weeks, I learned a vast amount of information about the space program, and in addition met many people in the aerospace industry - including engineers, physicists, computer technicians, economists, and even a few psychologists and sociologists.

A common thread of inquiry throughout the course of this program addressed the role of the public and the public's opinion about the space program. It was frequently assumed that public attitudes toward the space program are of considerable importance in the development of social policy relative to the space program. I was intrigued with the possibility of studying public attitudes toward the space program, while considering the union of this intrigue to my desire to find a research topic for a master's thesis. I called Rockwell International and talked to the Public Relations Director, informing him I was returning to Chicago, and would like to investigate public opinion about the space program in that area. I asked if he had any suggestions. He did. He informed me of Chicago Spacewatch, a community and educational program designed to stimulate public awareness and understanding of the space program, sponsored by the National Space Institute (NSI). He suggested I contact NSI under the guise of perhaps conducting a program evaluation

of this program. After several attempts to reach the Vice President of NSI, I finally made contact with the Director of Communications, Tom Gorski, who confirmed in a brief five minute conversation that NSI would be interested in an evaluation. Thus, I returned to Chicago with high expectations to conduct an exciting master's thesis.

To this day, many months after its completion, I am awed by the serendipity of the fact that I was in California to find out about Chicago Spacewatch. In my mind, it is doubtful that I would have heard of it while in Chicago.

Chicago Spacewatch was one month in duration, however, the evaluation took approximately seven months from acquisition to completion. I dealt with a number of problems, including political, technical, administrative and personal problems. Unstandardized treatments, a limited experimental design, a program essentially uncontrolled, inadequate program and research personnel, funding limitations, time constraints, and often a severe sense of inadequacy are some examples.

Everything I had ever read about program evaluation before I undertook this project and everything I have read since is much more real since the Chicago Spacewatch evaluation. It was indeed a learning experience, which is as it should have been. As a second year graduate student however, I was also faced with a year booked with courses, qualifying exams and a teaching assistantship. Without the help of Jill and John (my advisors) I truly wonder if I would have survived this tremendous learning experience. With such promising words as "good luck" and "get some sleep", I did begin one of the most valuable learning experiences in my career.

The ultimate reason in conducting this exciting, ferverous and at times helleous project, was a master's thesis. After submitting my final reports to NSI and closing the files on Chicago Spacewatch, I still had the task before me of preparing a master's thesis document. It was at this point that the redundancy of rewriting my reports into academic style did not seem to maintain the characteristic nature of the entire project - a learning experience. With suggestions from prestigious program evaluators in the field, such as Cook, Scriven, and Weiss, who state that metaevaluation (the evaluation of an evaluation) is a heuristic enterprise, I decided that an analysis of my own evaluation would be beneficial. In addition, since this was my first evaluation, and since I intend to proceed in this field, I felt the process of critically analyzing my own evaluation would be beneficial.

My master's thesis, therefore, is a complete, in-depth critical analysis of my program evaluation of Chicago Spacewatch. It is not an attempt to justify the mistakes I made, nor an attempt to heal any ego wounds. It is, as has been the entire project, a final chapter in the learning experience.

VITA

The author, Sherry R. McNeal, is the daughter of Earl James McNeal and Vivien G. Willcoxen of Minneapolis, Minnesota. She was born on December 6, 1948 in St. Paul, Minnesota.

Her elementary education was obtained in the public schools of St. Paul and Minneapolis, Minnesota, and secondary education at Edison High School in Minneapolis, where she graduated in June of 1966.

In the fall of 1966 she entered the University of Minnesota and throughout a course of ten years majored in English Literature, Classical Piano, Theater, Philosophy, Accounting, and Business; all of these on a part-time basis while working in the insurance industry. In 1977 she returned to school as a full time student at Northeastern Illinois University and obtained her Bachelor of Arts Degree in Psychology in June, 1978.

In September 1978, she entered Loyola University as a graduate student in the Applied Social Psychology Department, where she continues to pursue her education.

TABLE OF CONTENTS

LIST OF FIGURES

CONTENTS FOR APPENDICES

CHAPTER I

INTRODUCTION

Program evaluation has become a major field of social science research within the past decade. Stemming from roots in the fields of mental health and education, wherein evaluation became a tool to assess to what extent programs were providing intended benefits to program participants (Freeman, 1977; Patton, 1978a), the concept of evaluation has broadened its scope into several disciplines. It has become a major area of academic study and has become important in the development of social policies and management of social programs (Graycar, 1979; Freeman, 1977; House, 1976). Its universal application and potential benefits to any action program provide a variety of contexts for extensive utilization of social psychological techniques.

A. PROGRAM EVALUATION DEFINED

Due to the broadening scope of evaluation and the universality of its application, a diffused definition of evaluation has resulted (Patton, 1979). A global definition might include a judgement passing activity (Graycar, 1979), or any study which provides information which will reduce uncertainties. A more refined definition includes those studies which provide significant contributions to decision making processes (Weiss, 1972; Edwards, Guttentag and Snapper, 1975). Program evaluation may therefore include studies which provide descriptive information, are exploratory in nature, monitor program implementation, investigate social indicators, or produce information which is not utilized in any meaningful way.

To mitigate the problems of definition, several different types of evaluation have been identified. The most common dichotomy of program evaluation studies is that of formative and summative evaluations. Formative evaluation provides information which is fed back to program planners during the development of the program's curriculum to help improve it (Weiss, 1972; Patton, 1978a). More specifically, formative evaluation describes program operations, identifies effects produced by the program, determines the nature of the problems being addressed, and can be undertaken to observe the effects of different means of implementing the program in order to modify and develop the program (Rutman, 1977).

Summative evaluation is done after the curriculum is finished, thus providing information about the effectiveness of the program and subsequently providing input into the decisions of program continuation (Weiss, 1972; Patton, 1978a).

Likened to the schema of formative and summative evaluations is a categorization system of process evaluation and outcome evaluation (Freeman, 1977; Cain & Hollister, 1972). Process evaluation is a systematic assessment of whether the program operates in conformity to its design, reaching the specified target populations. Process evaluation therefore includes administrative monitoring. Outcome evaluation is a measurement of change toward the desired objectives with the possibility of a cost/benefit analysis. A comprehensive evaluation would be a combination of the two.

Similarly, Schulberg and Baker (1971) suggest two models of evaluation: the system model and the goal attainment model. The system model establishes the degree to which an organization realizes its goals

under a given set of conditions, thus considering such parameters as the effective coordination of organizational subunits, the acquisition and maintenance of necessary resources, and the adaptation of the organization to the environment and to its own internal demands. The goal attainment model measures the degree of success/failure in reaching predetermined objectives.

Another schema for classifying evaluation is to consider how the evaluation pertains to different facets of the program. These facets might include program planning and development, project monitoring, impact assessment, economic efficiency, or a combination of any or all of these (Rossi, Freeman and Wright, 1979). Different questions may be asked for each facet of the program, and thus an evaluation is cataloged according to the questions it answers.

Considering that an evaluation is tied to a decision making process, Alkin (1972) notes the development of a decision-oriented classification of the various types of evaluations, including: systems assessment, program planning, program implementation, program improvement, and program certification. Reicken (1972) suggests a schema based on the type of study to be conducted including: effect studies - the degree to which the program objectives are achieved; operations analysis - emphasis on the means or operations of the program without attention to the ends' surveys of need, assessing the need for or desirability of a contemplated action and investigation; and the independent audit, or policy type of evaluation. With respect to federally funded evaluations, Wholey, Scanlon, Duffy, Fukumoto and Vogt (1970) suggest that evaluations should be catalogued according to the scope of one's perspective, such as a national vs. local endeavor, or a

smaller project within a larger program itinerary.

The definition of evaluation and its subsequent label within a given classification schema are dependent upon several factors of importance depending upon one's perspective. These alternative perspectives are also evident in assessing the qualify or calibre of a given evaluation, a judging process referred to as metaevaluation, or an evaluation of an evaluation (Scriven, 1976; Cook and Gruder, 1979).

B. METAEVALUATION

According to Cook and Gruder (1979), the concept of metaevaluation is not new. It was first proposed in the educational literature in 1940 (Orata, 1940, as reported in Cook and Gruder, 1979). Scriven (1969) provided the term "metaevaluation" for the concept.

Metaevaluation has become increasingly important in the field of program metaevaluation for several reasons:

- Program evaluation is a research endeavor, subject to critique and review by peers and colleagues.

- Program evaluation provides data, results, conclusions and recommendations to program planners. If these data are not based on sound theoretical and empirical research practices, program planners may make decisions on faulty logic, and/or data. There is a responsibility to the client, the program planners.

- Particularly in summative evaluations, there is considerable power in evaluation data. In some cases, it may determine whether a program is to be continued, discontinued, or placed on probation, so to speak. Or, in other cases, data may be ignored if the results of the evaluation are not reflective of the expectations of the program planners; this action may be justified on the basis of a poor

evaluation. Evaluations have become a tool used throughout the role heirarchy in political game playing. Fear of survival is a major repercussion of the political interface of program evaluation and allottment of funds. The fact that so many evaluations have been negative has enhanced this fear. To offset the deprication of the program itself, program planners have attacked evaluation methodology and process.

- Evaluation data may provide ideas and information for proposals for reform - in which we may try out new programs designed to cure specific social problems. These programs may be retained, imitated, modified, or discarded on the basis of the program's apparent effectiveness, determined by an evaluation (Campbell, 1975).

- The Federal Government engages many evaluations to analyse the effects of public programs (Wholey, et al, 1970). The quality and calibre of evaluation data is often instrumental in making decisions of federally funded programs.

Although procedural models for conducting a metaevaluation are not plentiful, there are some guidelines available. Patton (1979) suggests that evaluation should be utilization focused; thus, the utilization of findings is an important criterion in evaluating the quality of an evaluation. In a decision making approach to evaluation, the quality of the evaluation is revealed in resultant changes to a program and the degree to which decision makers report the use of evaluation findings (Edwards et al, 1975). If impact assessment is the focus of the evaluation, then goal specification and methodological rigor are crucial. Scriven (1976) maintains that metaevaluation standard operating procedures should include a replication of the research

involved in the evaluation.

Cook and Gruder (1979) suggest models of metaevaluation research which can be carried out to improve the technical quality of empirical summative evaluations -- which are those "studies where the data are collected directly from participants within a systematic design framework" (pg. 470). The models proposed by Cook and Gruder are dependent upon three factors: 1) the time the metaevaluation takes place (i.e. during or after the evaluation); 2) whether the data are manipulated by the metaevaluator; and 3) the number of independent data sets that can be used to evaluate a particular program. For example, if the metaevaluation is being conducted subsequent to the primary evaluation, and the data are not manipulated, and there is a single data set, the model for a metaevaluation is an essay review of an evaluation report. If the data are manipulated, yet the other two factors are the same, the model would be an empirical re-evaluation of an evaluation or program. The essay review is the general approach used in the proposed metaevaluation herein.

C. A PROPOSED METAEVALUATION

The proposed research project is a metaevaluation of a program evaluation conducted by myself. The program evaluated was Chicago Spacewatch, a month-long program of community and educational programs intended to stimulate awareness and interest in the U.S. Space Program. The activities of Chicago Spacewatch included a variety of means to disseminate information about the space program to the people of Chicago, including general media, community programs and educational programs. Chicago Spacewatch was sponsored by the National Space Institute (NSI), the Chicago Public School System, the National

Aeronatics and Space Administration (NASA) the Museum of Science and Industry, and other organizations. The program was one month in duration; however, the evaluation took approximately seven months from acquisition to completion.

It should be noted that due to funding limitations and time constraints, metaevaluations in general are often impractical. Decisions need to be made and a secondary evaluation adds considerable time. To insure objectivity, a third party evaluator should be obtained; time and money are typically not allocated for this purpose. The proposed metaevaluation is not subject to these general limitations, however. The metaevaluation is an academic exercise. It is not being funded and the results will not be made available to the primary sponsor of the program evaluation. It is proposed as a heurestic, beneficial enterprise for the following reasons:

- a retrospective analysis of one's own evaluation can provide a better understanding of the rationale behind the evaluation methodology and process, the constraints and limitations of evaluation, and cogitations of alternatives perhaps not previously considered.
- investigating and defining possible mistakes in the process of the evaluation can be extremely useful in consideration of future evaluations to be conducted.
- particularly for the novice evaluator, this process would enhance training in the skills of conducting a "good" evaluation.

The method of evaluation reported herein is unique, developed exclusively for purposes of an evaluation of the Chicago Spacewatch Program Evaluation. The general approach, however, is an essay review, as suggested by Cook and Gruder (1979). Chapter II is a discussion of

the method; the remaining chapters are the results of the metaevaluation. Appendix A is an Executive Summary of the Program Evaluation of Chicago Spacewatch, which includes a description of the program, a discussion of the evaluation research methodology, results obtained and major conclusions drawn.

CHAPTER II

METHOD

The method of metaevaluation used to evaluate the Chicago Spacewatch Program Evaluation consists of an analysis wherein a comparison of an ideal program evaluation is made with the real program evaluation of Chicago Spacewatch. An essay review or discussion of each step in the evaluation will be analyzed on the basis of this comparison. The deviations of the real from the ideal will be discussed, as well as the reasons for the deviations (if any) and the problems encountered in attempting to make the real, ideal. This method, therefore is twofold. One is to determine the characteristics of the ideal program evaluation most appropriate for an evaluation like Chicago Spacewatch. The other is to draw comparisons between this formulated ideal and the real Chicago Spacewatch evaluation.

To establish the parameters and process of the ideal program evaluation, the major source of information is program evaluation literature, which is often replete with suggestions to attain an ideal. "The ideal" however, is dependent on the type of evaluation conducted, and to a certain extent the accepted definition of program evaluation. Thus, there is controversy regarding what the "ideal" is. If the controversy is relevant to the Chicago Spacewatch evaluation, both sides of the argument are presented and incorporated into consideration of the ideal program evaluation. This effort has resulted in a process model of an ideal program evaluation, which represents a synthesis of the evaluation literature deemed most appropriate.

The metaevaluation is conducted on the basis of a comparison of this established ideal and the real circumstances of the Chicago Spacewatch program evaluation. For each step in the process model, the limiting circumstances (e.g. time constraints and limited funding) will be identified, and the possible effect of these circumstances and other problems encountered during the course of the evaluation on the quality of the evaluation will be reviewed. In addition, the perspective of the program planner (as assumed by the evaluator) will be discussed where relevant. Similarities to the ideal will be noted, and deviations from the ideal will be discussed, reviewed and alternatives suggested.

Sources of information regarding the real program evaluation of Chicago Spacewatch, include:

- thoughts and feelings of the evaluator occurring during the analytic process and those recorded during the acutual evaluation.
- feedback about the evaluation from relevant others, including advisors, sponsors and other persons contacted during the course of the evaluation.

This analysis is subjective and retrospective. It is being conducted nearly two years following program implementation. The method is not precedented in the literature; however, it is likened to a metaevaluation model identified by Cook and Gruder (1979) in which an essay review of the major conclusions made in the evaluation is prepared subsequent to primary data collection. In addition, advisement and recommendations are presented on the basis of lessons learned.

CHAPTER III

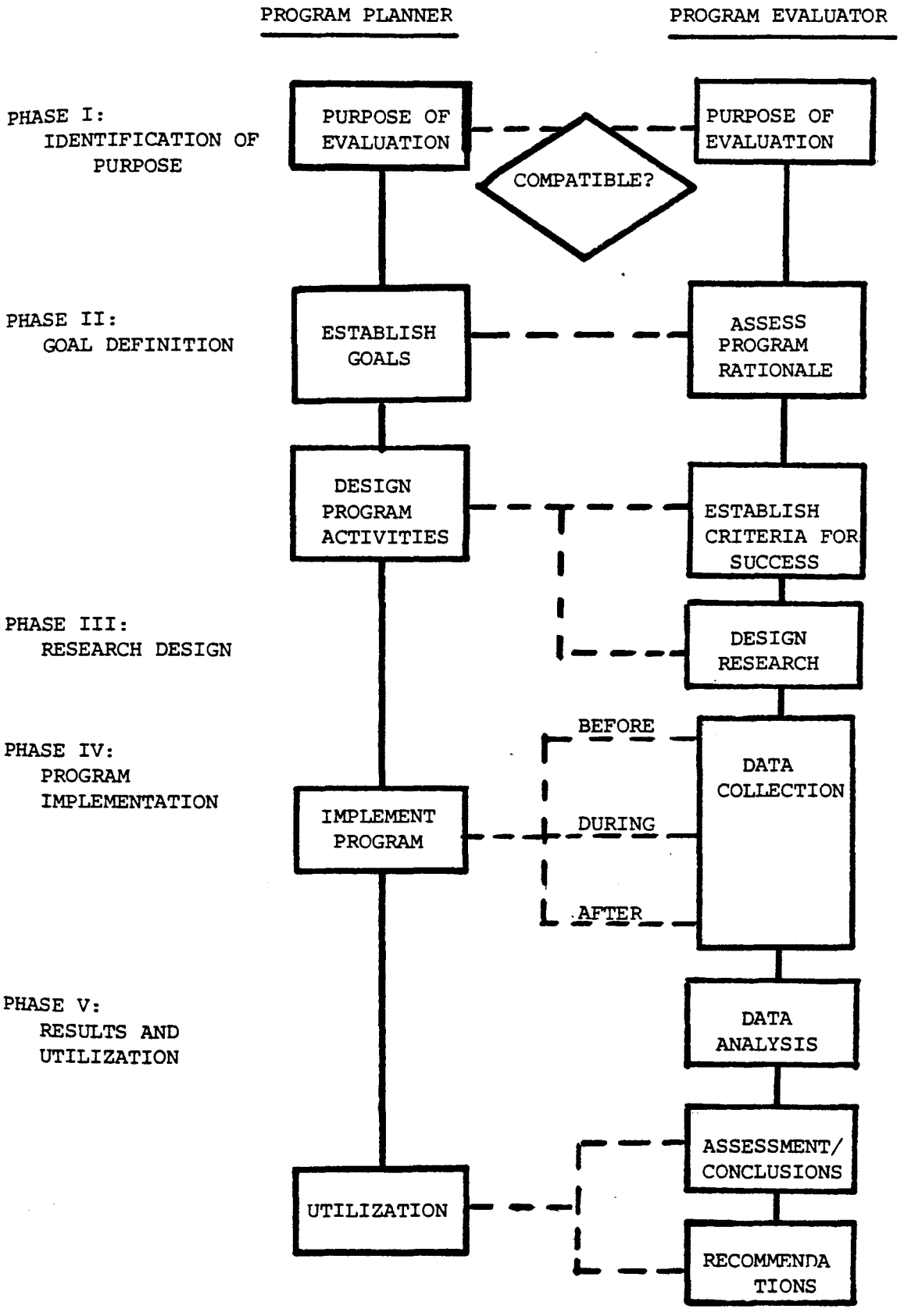
THE PROCESS MODEL OF PROGRAM EVALUATION

The major elements of the process model of program evaluation, symbolic of the ideal in program evaluation, are presented in Figure 1. It is divided into five phases which follow a linear time frame, beginning with an identification of the purpose of the evaluation, followed by a phase for goal definition, research design, program and research implementation, and finally results and utilization. Not all evaluations require an emphasis of each area; however, it is assumed that in the ideal circumstance, all areas will be addressed. The model, therefore, intends to incorporate most perspectives evident in program evaluation literature. The major features of the process model illustrated in Figure 1 are intended to represent a summary of the model. Each of the five phases is broken down into considerably greater detail in each of the following five sections of the report.

The model involves a series of activities which are linked and interconnected by a series of dark heavy lines, which suggest a stochastic process in which one cannot or should not proceed without satisfaction or completion of the preceding one. The dotted lines are indicative of a communication flow between program planners and program evaluators. Although specific tasks are allocated to each, this communication flow is suggestive of a joint effort to accomplish each respective task.

The model distinguishes between the planners and evaluators for several reasons. The "real" is incorporated into the "ideal", in that

FIGURE 1: SUMMARY CHART OF IDEAL PROCESS MODEL



there is likely to be differences between planners and evaluators. Twain (1975) suggests the following differences:

- Ideology of the researcher tends to be incongruent with that of program administrators.

- There are differences in career patterns. For the researcher, the evaluation may be a stepping stone in career development. The planner, however, is looking for credibility of existing techniques, rather than searching for new ones.

- The basic values and assumptions of the agency may be the basis for the agency status and recognition, and also the vehicles of action. The evaluator is trained to question these values and assumptions, thus engaging in a process which may immobilize the agency.

There are, therefore, some activities which by their nature many create conflict between the planner and evaluator. It is possible that program planners may not be concerned only with achieving program goals, but also building long-term support for the program (Weiss, 1975). The program can become a political tool, thereby marking a significant difference between the planners and evaluators. In addition, there are other sources of potential friction identified by Gurel (1975):

- identifying program objectives, rationale and procedures
- motivations for the evaluations
- demands of the operating staff.

These potential friction areas are important to identify and thus Phase I, an identification of the evaluation purpose, for both the planner and evaluator, is an important step in the ideal evaluation. If the purposes of planner and evaluator are not compatible, the evaluation should stop at this point.

The process model demonstrates considerable communication between planners and the evaluators; ideally, suggestive of good interpersonal relationships in the planner-evaluator interaction. Although it is important for the evaluator to be objective, there are also strong advocates of participatory evaluation research (Freeman, 1979). Ideally, the evaluator must be sufficiently objective to avoid a biased evaluation in favor or disfavor of the program planners, yet sufficiently involved to know and understand all motives and objectives of the program and purpose of the evaluation. According to Rossi and McLaughlin (1979) a fundamental aim of evaluators in interactions with administrators must be to obtain the information necessary to make the most effective choices in planning and carrying out the evaluation. This aim is considered very heavily in the process model, where continual interaction between the planners and evaluators suggests a participatory role of both, in each other's roles. Therefore, the planner has input into the evaluation research design, and the evaluator has input into program planning. This interaction is considered to be important throughout the entire program development and implementation.

CHAPTER IV

PHASE I: IDENTIFICATION OF PURPOSE

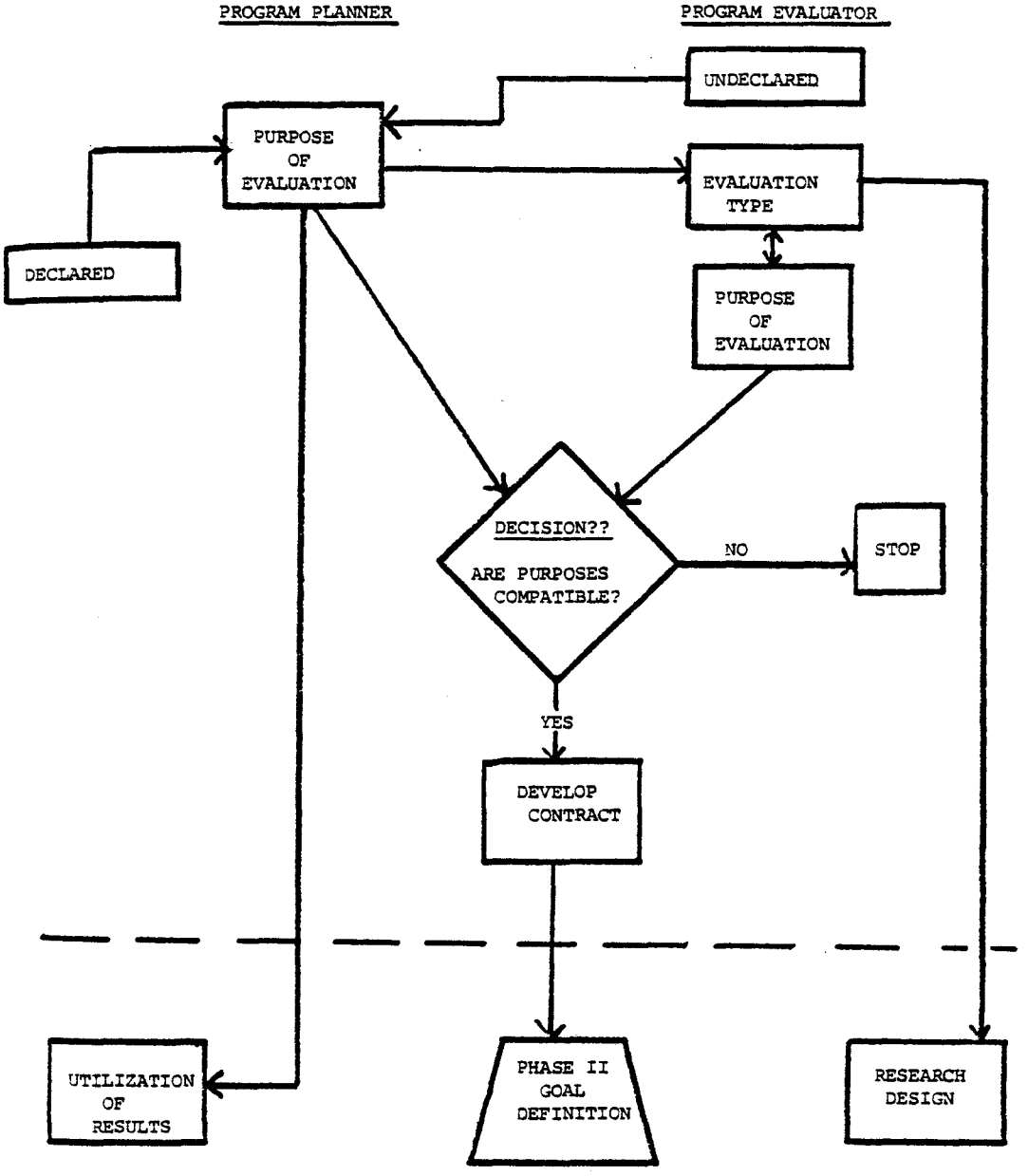
A. THE IDEAL

According to Weiss (1972), lesson number one for the evaluator newly arrived on the scene is: find out who initiated the idea of having an evaluation of the program and for what purposes. Who's asking questions about the program? Who's willing to pay for an evaluation to get answers? As identified in the process model of Phase I (Figure 2), the declared purpose by the program planner may or may not be the only or real purpose in mind. In fact, it may be impossible to identify all the reasons for an evaluation; however, it is important for the evaluator to identify as many as possible, including any that may be political.

It is suggested that evaluation is by nature a political activity (House, 1976), for program planners are not just concerned with achieving program goals, but also building long-term support for the program (Weiss, 1975). Thus, evaluation may serve decision makers as a tool in determining re-allocations of resources and legitimizing who gets what funds. There may also be hidden or undeclared reasons for the evaluation. An example might be what is referred to as a psuedo-evaluation (Rutman, 1977), in which the evaluation is used to trigger a decision, without regard for the data obtained in the evaluation. An "eyewash" is a pseudo-evaluation in which a deliberate focus on the surface appearance of a program results in a program which "looks good." Although the undeclared purposes of an evaluation may be

FIGURE 2

PROCESS MODEL OF PHASE 1: IDENTIFICATION OF PURPOSE



difficult to uncover, it is important for the potential evaluator to do so for serveral reasons:

1) As identified in the process model, the reasons given for an evaluation will determine the type of evaluation to be conducted (e.g. formative or summative). The type of evaluation is a major influential factor in selecting appropriate research technology.

2) The purpose of the evaluation provides considerable input into the goals of the program (Weiss, 1972; Rutman, 1977) as well as the probability of the results being utilized by program administrators (Schulberg & Baker, 1971).

3) To avoid political implications, Brickell (1978) advises finding out at this stage what the client has to gain or lose from the evaluation and how the data will be used, reassuring the client that findings can lead to useful suggestions.

The evaluator should also determine what his/her purpose is in conducting the evaluation. Twain (1975) suggests that there should be mutual payoffs for both the evaluator and the program planner, and that immediate and potential benefits for both must be recognized at the outset. As previously suggested, program planners and evaluators are operating from different perspectives and may therefore have a different orientation, resulting in differing purposes of the evaluation research under negotiation. Ideally, both evaluators and planners should be honest and open in stipulating their purposes for the evaluation.

At this point in the process model, a decision is made: Are the purposes of the planner and evaluator compatible? To some extent this decision making process is a preliminary assessment of the evaluator's potential contribution to the program. As Weiss (1972) has stated, an

evaluation is not worth doing in the following circumstances:

- When there are no questions about the program.

- When the program has no clear orientation.

- When people who should know cannot agree on what the program is trying to achieve.

- When there is not enough money or no staff sufficiently qualified to conduct the evaluation.

If any of these circumstances prevail or exist, it is unlikely that the evaluation will provide a meaningful contribution to the development of the program. If purposes for the evaluation cannot be identified, or if the purposes between planner and evaluator are not compatible, ideally the process stops. If compatibility seems assured, a contract is developed and Phase II - Goal Definition begins.

This decision can be made by either or both parties, for if the evaluator will not consent to do the type of evaluation desired, the planner may go elsewhere to find a more cooperative evaluator. Or the evaluator may decide that he/she does not wish to take the job. However, if the decision is that the purposes are compatible, a contract should be developed. Budgetary negotiations will be important at this time. Preliminary research design and budget proposals should be considered. Although the budgets may be subject to change and modification, parameters should be considered, for they will have a direct influence in the planning stages wherein the research design will be finalized.

B. THE REAL

If the lesson of the newly arrived evaluator is to find out who initiated the idea of having an evaluation of the program, and for what

purposes, it would seem that the program evaluation of Chicago Spacewatch was in trouble from the start. Since I initiated the evaluation, it is probable that NSI would not have independently sought out a program evaluator. In a sense, NSI was"sold" the idea of conducting an evaluation. NSI's purposes for the evaluation are not clear; however, my own personal reasons were made clear to NSI from the onset.

1. Personal Reasons

Conducting a program evaluation of Chicago Spacewatch served my needs in several ways:

- it provided a topic for a master's thesis
- it involved research related to the space program
- program evaluation was a perferred research endeavor.

In addition, I had hoped to provide valuable information for the development of future programs related to the space program.

I had only to convince the representatives of NSI that a program evaluation would be a worthwhile project. When I finally reached my contact (Gorski) by phone, he was sufficiently interested to establish a meeting to discuss the possibilities of an evaluation during his next trip to Chicago. In preparation for the meeting, I developed several possible experimental scenarios to evaluate the program, including possible funding needs.

2. NSI Reasons

My approach in selling the program evaluation was to convince Gorski that a program evaluation would be beneficial. Evaluation data can provide a substantial source of credibility. Other Spacewatch activities were being planned for the future, and evaluation data

exemplifying how and why a given Chicago Spacewatch program might work were obviously of great value in future planning. In addition, an important variable in negotiations was the fact that I was offering NSI a very financially attractive deal. The proposal did not include any funds for my time.

In retrospect, it seems Gorski was prepared to proceed before our meeting, in that it didn't seem to take much effort to sell the idea of an evaluation to him. We agreed that I should prepare a proposal, with two or three possible scenarios and the cost of each.

The proposal stated specific reasons why an evaluation would be beneficial, using these reasons as stated objectives of the Chicago Spacewatch evaluation. The following excerpt from the proposal illustrates:

"An evaluation has been found to be a very productive adjunct to any program design. In particular it can:

- establish a clear and specific criteria for success
- provide a judging process that is more accurate and objective than intuitive evaluations.
- collect evidence of the program's effectiveness for a representative sample of participants.
- provide data that reduce uncertainties and clarify gains and losses (objective criteria for cost/benfit analysis).
- draw conclusions of effectiveness, merit and success.
- provide objective input in future decision making (i.e. what activities should be continued, discontinued, expanded, etc.)

The above objectives are the objectives of the Chicago Spacewatch program evaluation proposed herein. Although there are several potential levels of an evaluation, and several alternatives dependent upon the degree of sophistication, the enclosed evaluation has been selected on the basis of time allotment and economy. It is projected to provide those minimum measures which are necessary to produce a general but powerful evaluation of program effectiveness."

After briefly outlining possible experimental procedures, I concluded by proposing that my evaluation would produce:

- indicators of goal achievements and program effectiveness
- input for future decision making
- specifics for future funding considerations
- instruments tested for reliability, reuseable in future programs
- base-line data of space awareness and space attitude for use in future city comparisons, as well as in a comparative analysis of program effectiveness.
- a unique stimulant, in that the process of evaluation in itself, is an awareness generating technique.

Unfortunately, it is impossible to discuss NSI's purposes for the evaluation, in that I did not attempt to identify them at any time. In retrospect, however, it is possible to make some assumptions. It seems that Gorski was looking for some evidence which would document the success of Chicago Spacewatch and thus facilitate acquisition of additional funds for future projects. In addition, I believe there were excess funds in the Chicago spacewatch program budget, and I provided a meaningful way to spend some of that excess.

Any political or hidden motives for sponsoring the evaluation that NSI may have had are pure conjecture on my part. However, I do believe they existed. Unbeknownst to me at that time, NSI was experiencing considerable difficulty with related organizations such as NASA and aerospace companies. The Board of Directors for NSI is comprised of individuals from several such organizations, and debates were ensuing regarding the acquisition and use of funds to support the concept of NSI - to educate the public on the benefits of the space program.

Gorski, the man with whom I had immediate and direct contact throughout the evaluation had intended to leave NSI as soon as Chicago Spacewatch was finished, which he did. He knew of his intentions to leave the organization prior to our first meeting, although this information was not disclosed to me until after my final reports were submitted, and then somewhat by accident. It is rumored that the Vice President of NSI was later asked to resign due to disagreements on the distribution of funds and the internal structure of the organization. This resignation was subsequently rescinded.

It seems unlikely that these political upheavals which were brewing in the NSI organization had more than an indirect effect on my evaluation. It is my belief, however, that both Gorski and the Vice President were looking for support for the program, and posssibly hoped that the evaluation would provide impressive data which would demonstrate the appearance of a "good program" (possibly an eyewash or pseudo-evaluation). I was totally unaware of these circumstances during the evaluation and at no time did I heed the advice suggested by Brickell (1978) - to find out how the data would be used or to find out what the client had to gain or lose from the evaluation. My naive belief at that time was that NSI was adopting my proposed reasons for the evaluation and that they had good intentions of using the evaluation data in program planning of future Spacewatch activities.

3. Lessons Learned

In the ideal process model, an identification of the purpose for the evaluation by both evaluator and planner is critical. It provides the evaluator with necessary information to provide an evaluation which answers the pertinent questions posed by the program planner. It also

gives the evaluator some indications of how the evaluation results will be used. Although my purpose in conducting an evaluation of Chicago Spacewatch was stated and understood, the NSI purpose for the evaluation was never identified; it was assumed. Lack of utilization of evaluation results (discussed in Phase V, Chapter VIII) is due in large part to this fact.

The evaluation provided too little descriptive information (e.g. number of people attending the Museum on a given program day) and too much statistical analysis. The evaluation did not conform to a desired time frame by NSI. This later became evident, when I discovered that data were needed for a Board of Directors's meeting in January following the program. I was still collecting data at this time. Had I been aware of the time table, I might have altered the design to provide more return on data more quickly.

The potential value of seeking out the purpose of the evaluation is therefore the major lesson learned. If the ideal process model had been followed, it is conceivable that the evaluation may not have been done or that it may have changed substantially. Although I had designed the evaluation on the basis of stated goals of the program, the evaluation may have been designed to better fit the needs of NSI. It is therefore important to determine at the beginning what the purpose of the evaluation is, how the data will be used, when data are needed, etc. Answers to these questions will provide guidelines in the development of the evaluation research design.

CHAPTER V

PHASE II: GOAL DEFINITION

A. THE IDEAL

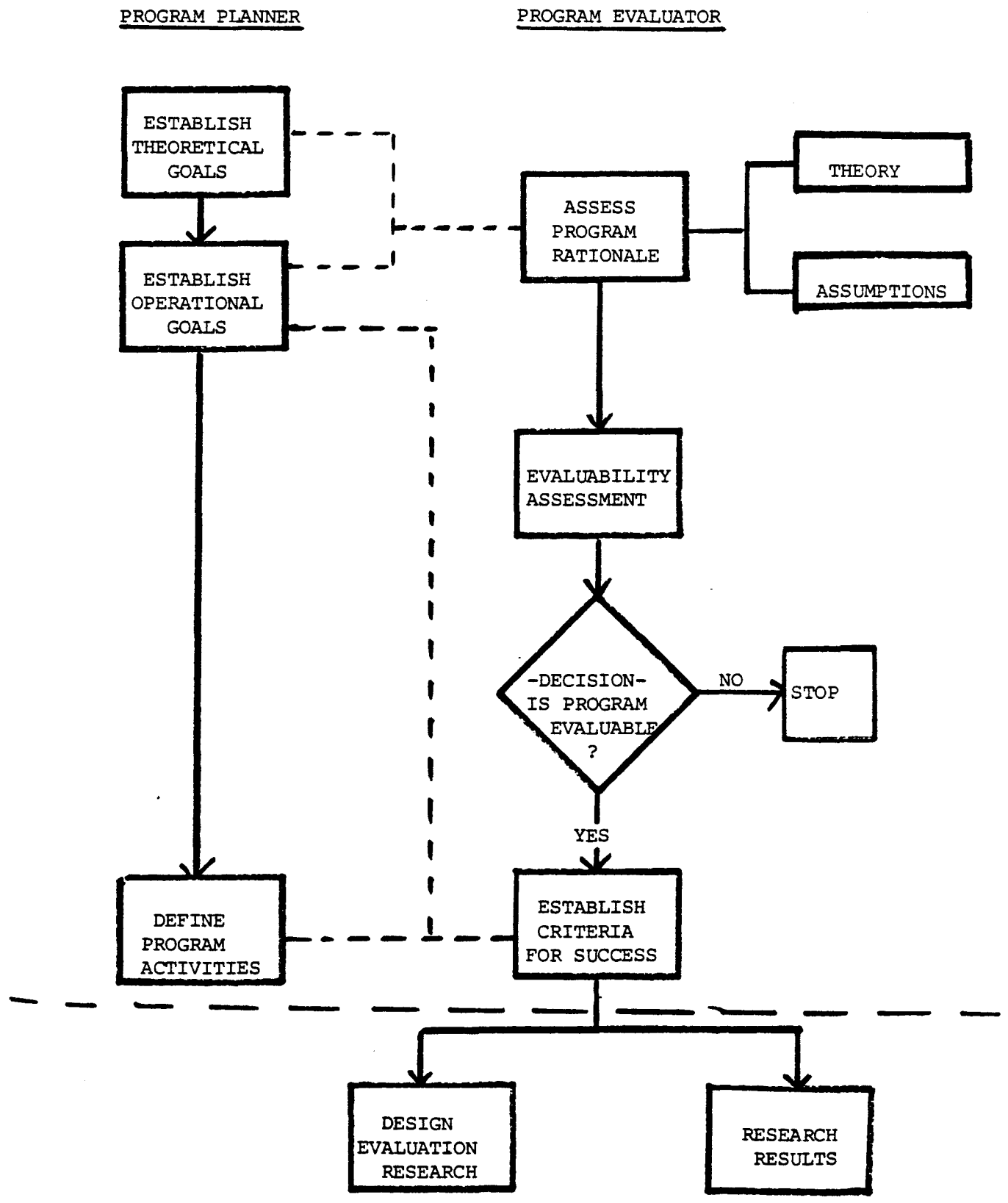
According to Weiss (1972) the traditional formulation of the evaluation question is: To what extent is the program succeeding in reaching its goals? To answer the question, it is necessary to identify the program goals, translate the goals into measurable indicators of goal achievement, collect data on the indicators and lastly, compare the data with the goal criteria. In theory, from a detached perspective, the procedure seems simple enough; however, in the real situation, each of these steps becomes a major process onto itself. The identification of program goals and the translation of these goals into measurable indicators of goal achievement are indicative of only a part of the entire planning process which should, ideally, take place at this phase of the evaluation. The Process Model of Phase II: Goal Definition, is presented in Figure 3.

This phase of the evaluation process is perhaps one of the most critical, for the ultimate end of this phase is to define the program activities and design the research program. There is a need for considerable communication between planners and evaluators, and although the model suggests specific duties or tasks for each; ideally, both planners and evaluators should proceed through this phase hand in hand.

Ideally, the process model suggests for planners, a progression from the establishment of theoretical goals to operational goals of the program. The differentiation between theoretical and operational goals

FIGURE 3

PROCESS MODEL OF PHASE II: GOAL DEFINITION



is both quantitative and qualitative in nature. Theoretical goals are the unreal, abstract and often ambiguous goals that resound with promise, but carry very little validity. For example: improve education; rehabilitate the deviant; change people's attitudes; increase sales. The operational goals narrow the theoretical goals down to more quantifiable terms, such as increasing the number of classes available in the college curriculum by 10%; providing inmates of correctional institutions with training programs derived from major business entities, etc. The role of the evaluator is to assess if the linkage of these goals is theoretically sound on the basis of social psychological research and theory, or other relevant theory and research. Assumptions are frequently made by program planners which should be identified and assessed in terms of their soundness and/or significance. In essence, the role of the evaluator at this point is to assess the rationale of the program.

Although adequate theory is difficult to ascertain (Twain, 1975), it is still essential to address underlying assumptions and the nature and cause of the program. The theoretical linkage between outcome and treatment is particularly important and should be established beforehand (Boruch & Gomez, 1979; Larkey, 1979; Weiss, 1975; Twain, 1975; Nunnally & Wilson, 1975; Cohen & Weiss, 1978; Sechrest, 1979).

According to Sechrest (1979) theory has a critically important role in evaluation research, for he claims that it is crucial for the evaluator to be able to specify the theoretical link between the intervention planned and the outcome expected. The major reason for this is that integration of theory, conceptual analysis and methodology, provide the strongest possible foundation. The strength of the

treatment can therefore be evaluated in terms of theoretical premises supporting the use of the treatment, as well as the methods of measurement and the procedures of implementation.

For many programs, social science knowledge and theory would suggest that the goals are not well reasoned, that problem diagnosis and selection of interventions are inappropriate, and the chances of success are slight (Weiss, 1975). Thus, an understanding and/or investigation of theory can assist in placing the goals in a sensible perspective, resulting in more modest and realistic expectations.

Nunnally & Wilson (1975) point out that one does not measure objectives. One measures the attributes of objectives, and thus, measurement requires a process of abstraction. It is important, therefore, to carefully consider the nature of the attribute before attempting to measure it; it is possible that the attribute does not exist. For example, consider the measurement of attitudes and the presumed role of attitudes in behavior change. It is possible that the act of measurement results in creating the attribute; in that measuring an attitude can create an attitude. It may also be the case in which a change in attitudes is assumed to lead to a change in behavior; from a theoretical perspective, this assumption is questionable.

Assessing the rationale of the program, therefore, includes the understanding and identification of any theory behind the proposed program and the identification of the assumptions. In some cases, it may be that the program rationale is not sound, and thus the evaluator will discern that the program is not evaluable on the basis of the goals and proposed action to meet those goals. It may be beneficial for the evaluator to diagram the cause and effect linkages that lead,

hypothetically, from program inputs to desired end-states (Rossi & McLaughlin, 1979). This diagram may assist in communicating the importance of developing a sound rationale in the development of goals, both theoretical and operational. Since many programs are evaluated to assist the decision maker regarding the future of the program (Edwards, et al, 1975), this is the point where problems should be addressed, clarified and hopefully solved in goal clarification.

Formulating program goals is probably one of the most difficult tasks in the evaluation process, and one that is frequently side-stepped, ignored and inundated with a number of problems. The following are major potential problems:

- Goals may be vague, ambiguous, too general, fuzzy, and thus, immeasurable (Weiss, 1972; Twain, 1975; Patton, 1978).

- Conflicting goals may exist (Tabor, 1978). For example, there may be conflicts between administrators or members of the program staff - each with different ideas about what the goals of the program are (Twain, 1975; Weiss, 1972; Patton, 1978).

- Some goals are illusionary, in that there are not real intentions to attain these goals. They are used for window dressing - to make the program look good.

- There may be latent goals which may arise during program implementation, not previously considered, due to unanticipated consequences. These goals may usurp priority over those previously identified.

- The goals may be theoretically inappropriate, as previously discussed in terms of the theoretical rational of the program. It is frequently the case that goals are not well reasoned, indicating that

the problem diagnosis and selection of time and type of intervention may be inappropriate.

- The goals may be in a constant flux of change in order to meet the demands of officials, finances, and other administrative elements. At some point goals need to stabilize in order to adequately design the program activities.

- There may be undetected, unintended consequences of the goals which require change in mid stream (Rossi, Freeman, Wright, 1979).

- Goals may be inter-related to such an extent that it is not possible to change one without changing all the others (Schulberg & Baker, 1971). Because many of these problems are not infrequent, a considerable amount of evaluation literature discusses guidelines and recommendations for mitigating these problems. Weiss (1972) suggests that if the goals are not clear, specific and measurable, the evaluator should work with planners to obtain these objectives before developing a research design. Thus, the process model (again reflecting the ideal) is indicative of communication between planners and evaluators in the goal definition phase. Unfortunately, planners can become defensive when they are asked to defend and/or articulate the objectives of the program. It can be a touchy situation and the evaluator may need to employ considerable diplomacy and tact in this mutual undertaking of goal definition.

Rutman (1977) and Patton (1978) suggest that too much attention is focused on the goals. Rutman (1977) suggests that an exclusive focus on goals ignores the side-effects which may be more relevant. There are so many problems with goal definition <u>a priori</u>, that is may be more pertinent to analyze problem definition during the evaluation. In other

words, evaluation study would test the presumed connections, rather than attempting to establish the connections and potentially measuring inappropriate attributes of the program.

Patton (1978) suggests that in many circumstances goal clarification results in conflicts and political gaming. For example, the "fuzziness gambit" is a play in the goals warfare to deliberately establish fuzzy goals as a protective measure. An alternative is a goal free evaluation or a utilization-focused evaluation which provides information and data to program planners and decisionmakers regarding specific evaluation questions. The relevance of various information options to decisionmakers and information users in a particular situation can then be established. The evaluation is then channeled through to the ultimate end use - the utilization of findings. Unfortunately, with too much attention on the goals, and the outcomes as they relate to the goals, there is too little investigation into the utilization of findings.

In the ideal situation, planners and evaluators can cooperatively negotiate a set of goals that are theoretically sound, linking the proposed program with anticipated effects. The ideal situation is not always attainable, however, and it is at this point that the evaluator can profit from the answer to the question, "Is the program evaluable?" Several criteria might be used to reach a decision, including:

- the information gained from the rationale assessment
- the type of program
- the degree to which planners have adequately defined goals
- the degree to which cooperation between planners and evaluators has resulted in adequately defined goals.

If the answer to the question is "no", the evaluation process should ideally stop at this point unless major changes in the criteria to reach this decision are made. Obviously, if the evaluator does not feel the program is evaluable, it should not be evaluated. If the answer is "yes", the process continues. The evaluability assessment may also be based on several subjective, non-specifiable criteria. For example, just being able to get along with program planners is important.

The next step for program planners is to specify program activities (assuming a formative evaluation in which the activities are not yet implemented, nor totally defined). These program activities should be derived from the goals of the program, demonstrating a strong theoretical linkage between the goals and activities. Changes in the program, which may be considered as a result of the goal clarification process, should be made. Then, the evaluator, with the cooperation of the planner, should begin to define the criteria for success to be used in the evaluation research methodology. Conceptually, the evaluative criteria represent the basis upon which decisions are made about means toward ends (Suchman, 1972). The task is one of finding reliable and valid operational indices for measuring the attainment of some objective. The emphasis is usually technical, rather than conceptual. Both planners and evaluators should work toward a definition of success. It should be noted, however, that frequently the planner is overly optomistic, and thus unrealistic standards or criteria can develop. The evaluator has to maintain the objective role at this point.

Without clear, unambiguous, straight-forward goals, the task of establishing criteria for success can be nearly an impossible project.

And if criteria for measuring outcomes of success are undetermined, success is left totally undefined (Tabor, 1978), resulting in a very subjective evaluation of outcome data, subject to interpretation. Without valid outcome criteria, any changes produced or reasons why the program succeeds or fails will not be discernable, and thus, the evaluation will be a wasted effort.

For these reasons the evaluability assessment, a decision point for the evaluator, is included in the process model prior to the attempt to establish criteria for success. In essence, it is a reflection of the foregoing process of the goals clarification process and the assessment of the rationale of the program. It should be evident at this point whether criteria for success can be established.

Ideally, the results of this phase include a clearly articulated program, clearly specified goals and/or anticipated effects, a good rational linking the program to the goals and criteria for success. The evaluator can then proceed to design a research program to collect data which will provide the most meaningful information relevant to the goals. The results of the research will ultimately be compared to the criteria for success as a measure of program effectiveness.

B. THE REAL

The foregoing discussion of problems frequently encountered in evaluation research are all relevant to the evaluation of Chicago Spacewatch. The ideal flow of communication back and forth between administrators and evaluators during this phase, was virtually non-existent. The time frame of the acquisition of the evaluation will perhaps explain why this is so. I had previously received a brochure about Chicago Spacewatch from the Director of Public Relations at

Rockwell International, which outlined program goals and program activities. This brochure had been prepared several weeks prior to the actual program, and relatively few changes were made between the time the brochure was prepared and my first meeting with Gorski. The information presented in this brochure was the only information I had throughout the entire evaluation. Although Gorski intended to keep me up to date regarding any program changes, the only information I had with respect to program changes was obtained from my own observations and independent methods of acquisition.

The month long program of Chicago Spacewatch was set to begin October 1, 1978. The meeting with Gorski occurred in late August. I received approval on my proposal September 1st. I had less than one month to pilot test my instruments, and collect base-line pre-program data. The timing was so stringent that I barely had time to think about the goals of the program, let alone attempt to negotiate any changes. The only time I devoted to consideration of goals occurred during the one week when I prepared by proposal for NSI.

The Chicago Spacewatch Program Evaluation therefore began with an already specified goal program. There was no time to discuss, clarify, negotiate or change goals or activities. The pre-planning stage of evalution in the ideal format was thus non-existent, suppressed by time constraints. However, an analysis of the goals provided by NSI are appropriate as well as a post-hoc assessment of the rationale of the program for purposes of evaluating the Chicago Spacewatch Evaluation.

1. Theory and Assumptions

The overall program goal of Chicago Spacewatch was to stimulate public awareness, interest and understanding of space and the

application of space technology as potential tools in solving problems here on Earth. The fundamental assumption related to this goal is that an increase in awareness and interest in the space program will lead to attention toward information available about the space program, leading to a greater understanding of the benefits derived from the space program, and thus, ultimately, lead to a more pro-space attitude. In order to test the assumption that an increase in awareness and interest of the space program would ultimately lead to a more pro-space attitude, a measuring instrument was developed for the evaluation which measured knowledge of the space program, perceived benefits of the space program, and an attitude toward the space program. The instrument was intended to measure the correlation between variables, as well as discern any change in attitude before and after Chicago Spacewatch, possibly resulting as a consequence of increased knowledge and perceived benefits of the space program.

There is some evidence that public attitudes toward the space program, technology and other technologically oriented programs, are very highly correlated with the public's awareness of information pertinent to the program and perceptions of benefits resulting from the program. A study done by NASA in 1972 and 1974 (LaParte & Metlay, 1975) measured attitudes of the general California public regarding scientific research, technology and specific technological applications, such as those of the space program. The general consensus of this study was that public support for future technolgoy rests heavily on people's estimate of the probability of benefits and harms resulting from the implementation of the technology. The public does make a distinction between science and technology, perceiving scientific activities as

intrinsically beneficial, whereas perceiving technology as possibly leading to threatening outcomes.

The space program was one of several technological activities about which a more positive than negative attitude was found. In 1972, 60.8% indicated that the space program makes life slightly better to much better, versus is harmful, detrimental or threatening. In 1974, this percentage increased to 65.2% In particular, the benefits perceived from space travel included advancement in science and technology, and a relief of population pressures. The harms included cutting funds elsewhere, too costly, too dangerous, and the idea that God didn't mean us to.

Another study conducted by Opinion Research Corporation, funded by NASA (as reported in Laparte & Metlay, 1975) revealed that the public's attitude toward the space program is a function of the number of benefits perceived as resulting from the space program, in that the more benefits associated with the space program, the more pro-space attitude indicated.

The results of these studies provide evidence that knowledge of the space program and benefits resulting from the space program are positively correlated with an attitude toward the space program and space technology. These correlations were also found in the Chicago Spacewatch data. However, it should be noted that a correlation between these variables is not indicative of any causal relationship, as exemplified in the fundamental assumption of Chicago Spacewatch. It may be that a positive attitude towards the space program exists first, in which case an individual with a positive attitude may attend to information related to the space program and thus learn more of its

benefits. As this relates to the overall program goal of Chicago Spacewatch, it is possible that only those people already interested in the space program would attend to the various programs intended to increase public awareness. In other words, NSI would be singing to the choir, reaching only pro-space enthusiasts. The evaluation data support this proposition - in that those people attending many Chicago Spacewatch programs (such as the O'Neill lecture) had a very positive attitude toward the space program; yet they also scored high on knowledge factors and perceived considerable benefits from the space program.

Although the goals clarification phase of Chicago Spacewatch was virtually non-existent, had a discussion of this fundamental assumption occurred prior to program activity definition, several changes to the program may have resulted. For example:

- NSI was disseminating a lot of disconnected information about the space program which was not effectively linked together in a meaningful way. Educating the public about the benefits of the space program at a time when few people were even aware that a space program existed, was probably not the best approach. A greater declaration of the activities of the space program may have been more appropriate; in other words, letting people know that a space program was still in effect - preparing the shuttle and developing plans for a space operations center. A differentiation between knowledge of the space program, and benefits derived from the space program may also have been appropriate. A target of one or the other may have been more beneficial, rather than dispersed information related to each. Chicago Spacewatch itself had no focus, and thus, the public or program

recipients, did not focus either.

• Program orientation to particular groups of people, versus a general public audience may have been considered. Special exhibits at the museum and planetarium could have had a focus on a few specially chosen activities aimed at space interested people, rather than several public activities which were not well attended. Similarly with educators, rather than attempting to reach teachers in all disciplines, a better approach may have been to reach out to science teachers exculsively.

2. Goals and Program Activities

As the overall program goals was stated, the objective of Chicago Spacewatch was to stimulate awareness and interest in the space program. Ambiguity of this goal is evident from two perspectives. First of all, it is unclear how people would demonstrate awareness and interest in the space program. Would they attend special space related programs? Subscribe to space related journals? Talk to their friends about space? Secondly, if awareness and interest in the space program leads to a more pro-space attitude, it is unclear what behavior should be manifest to reveal or reflect a positive attitude. Would it be reasonable to expect individuals to write letters to congressmen requesting greater appropriations for the space program? Would Gallop polls evidence greater public support for space technology?

Unfortunately, anticipated consequences of Chicago Spacewatch, or expectations resulting from the effects of Chicago Spacewatch programs were not stated; nor is there any evidence to suggest that any thought was given to a definition of success. Criteria for success, therefore, were nearly impossible to establish. This is true not only of the

overall program goal, but also of more specific goals related to community programs (See Appendix A, pg. 14) and educational programs (See Appendix A, pg. 19). In only two circumstances was I able to identify criteria for success. One instance related to attendance at museum activities; the objective was to increase attendance to the museum over the attendance on record during the same time of year for several previous years. Another instance related to attendance to the teacher workshop held at the planetarium. The goal was to have a least 100 teachers attending. Although these two instances provided a measurable outcome, the outcomes were indicative of the success of program implementation, not the effects of the programs as related to the program goals. In other words, these outcome data did not provide any indications of impact.

In terms of the process model of evaluation, there is a big jump from the theoretical goals of the program to a definition of program activities. The crucial intermediary step of operationally defining these goals, which ultimately would lead to a discussion of the critiera for success, is void in Chicago Spacewatch. The goals were ambiguous and thus, success was nearly impossible to define. As a consequence, the program evaluation of Chicago Spacewatch also suffered. I had difficulty determining successes and failures. I found myself searching, collecting data here and there, attempting to find clear, objective indicators of impact. These indicators should have been established prior to the evaluation, not sought after during the evalution. As a consequence, data were collected that did not fulfill the needs of NSI nor provide relevant information for decision making. For example, I collected considerable data related to knowledge,

perceived benefits, and attitude toward the space program, which I, the evaluator, thought were important. However, these data were not particularly impressive to NSI personnel. It was later discovered that very little of these data were utilized. In part, this may be due to the fact that the all-purpose questionnaire did not address all aspects of the Chicago Spacewatch programs. There was so much information being disseminated, it was impossible to do so. Thus, it was not clear what information was being learned, what beliefs had changed, if attitudes were changing as a result of increased awareness and interest, or whether Chicago Spacewatch was in fact having its intended effect.

The goal clarification process is extremely important in program evaluation, and the lack of this process in the program evaluation of Chicago Spacewatch is evident in the lack of utilization of results, and the difficulty in measuring success of the program. Although it is not possible to determine how the program activities and/or evaluation may have changed as a consequence of a goals clarification process, it is probable that several improvements may have been made in identifying intended effects of the program, and the development of criteria for success for use in evaluating the program. Less data may have been collected, greater utilization of data may have occurred, and in general, a better program and a better evaluation may have resulted.

CHAPTER VI

PHASE III: RESEARCH DESIGN

A. THE IDEAL

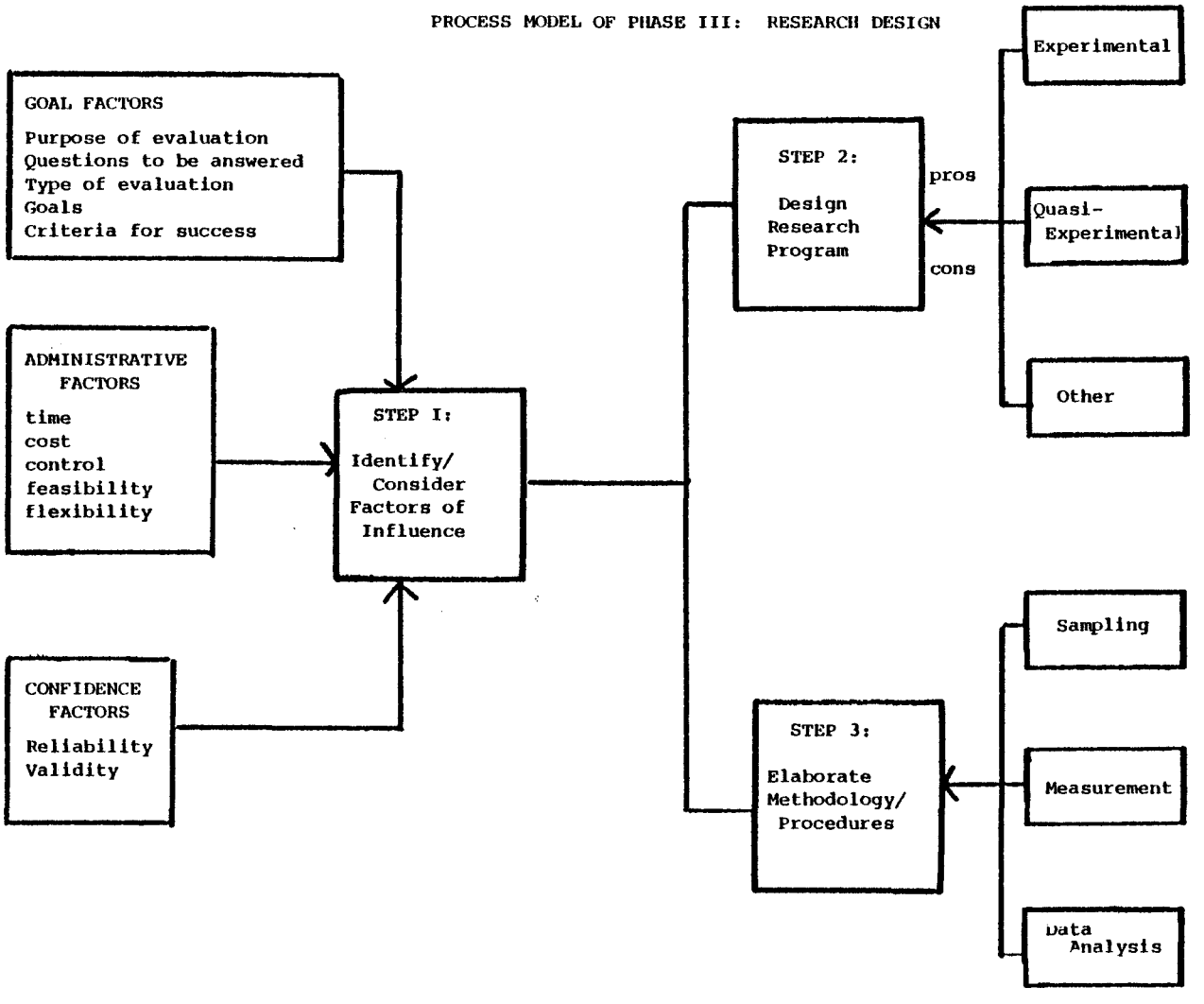
Developing a research program is the objective of Phase III, and it addresses most directly the research facets of program evaluation. The steps involved in this phase are critical, for the research design is subject to review and criticism following the evaluation. The research design determines to what extent the data collected are meaningful, valid, reliable, appropriate for the questions attempted to be answered, etc. All other phases of the evaluation are directly related to this phase. The preceding phases provide direct and meaningful input into the development of the design; the impact is significant in following phases, particularly the utilization of results by program planners.

The process model for Phase III is presented in Figure 4. Although this is primarily an activity for the program evaluator, ideally, both program planners and evaluators work cooperatively through this phase. As indicated in the model, there are three basic steps.

The first step is an identification and consideration of factors of influence in designing a research program, including those factors related to the goals of the program, administrative factors, and factors indicative of a level of confidence in one's conclusions and results of the research. In the second step, the development of a research program, there are a number of choices and decisions to be made, including decisions regarding the type of research to be conducted:

FIGURE 4

PROCESS MODEL OF PHASE III: RESEARCH DESIGN



experimental, quasi-experimental or other research that is not either of the other two. Step 3, the elaboration of the methodology, includes the specification of procedures related to the overall research design, such as sampling procedures, selection of dependent variables, measurement techniques, instrument design, and the proposed plan for data analysis. In the ideal circumstance, the factors of influence should be identified, weighted and considered with respect to both the type of research or research design, and the methodology. Trade-offs are inevitable, yet ideally in joint consideration, both evaluator and planner can determine the best choices and make decisions most appropriate for the program. In the real situation, there are a number of problems and controversies associated with this phase.

1. Factors of Influence

a. Goal Factors

The goal related factors of influence include the determinants of the preceding phases:

- questions of the program planner
- purpose of the evaluation
- type of evaluation best suited to meet the purpose (i.e. formative vs. summative)
- goals of the program
- criteria for success.

These factors will determine to a large extent the type of measures necessary to provide appropriate data. Evaluation research, in the ideal sense, is a developmental process in which each successive step is dictated largely by the knowledge gained in previous stages and related to the whole theory and conceptualization (Twain, 1975). Thus, the

preceding steps (phases) are major factors of influence in the design development.

b. Administrative Factors

Administrative factors of influence are those which may be to a large extent out of the control of the evaluator. However, they should be investigated, identified and considered in the ideal evaluation. These administrative factors include:

1. Time

One of the difficulties in any research is knowing in advance how much of a time lag exists between the implementation of a treatment and the manisfestation of an effect (Bernstein, Borhnstedt and Borgatta, 1976). It is possible that the effects may be gradual, continuous, or may occur all at once. When to collect data, and the frequency of data collection is administratively relevant in terms of cost. It is also important that program planners understand this concept, for if there is an insistance that measurements be taken at inappropriate times (too soon or too late) the evaluation researchers may arrive at incorrect conclusions about the effects of the treatment (Bernstein, et al, 1976). The time schedule, therefore, must be looked at and carefully examined to insure that there will be enough time for proposed research (Rein & White, 1978). In many cases, time constraints will prohibit the more desirable designs. A good design is useless if it does not fall within the program schedule. In general it is important that all appropriate measures can be taken in the time expected and/or allotted.

2. Cost

The budget is an important factor for consideration

because in the real situation it is possible that the best plans may not be affordable. One must design a research program within the limits of the budget. Since a better program is always possible, the ideal strategy should be one of developing the best research program within the affordable limits.

3. Control

The factor of control relates to the degree of experimentor control over the research. In the ideal situation the staff and evaluators are cooperatively working together. However, in the real situation, this is not always the case; yet cooperation from the staff may be crucial if elements of the research design require an invasion into staff privacy. Considering the character of the people one is working with is therefore an important factor in developing the research design (Freeman, 1977). It is also possible (in the real situation) that program and project operators will not be cooperative in permitting evaluators to modify projects in terms of participant selection, treatment variation, etc. (Williams, 1971). Even if these modifications are allowed, it remains questionable whether participant selection procedures and design modifications will be implemented properly or carried through for a sufficiently long time to permit a meaningful evaluation. With an on-going program, it is probable that program staff have developed ways of doing things that are not easily changed, adjusted or controlled to allow for more experimentor control. There may be a number of administrative procedures - not easily detected at first glance - which may ultimately interfer or disrupt experimentor arrangements. Cooperation between program staff and evaluators can therefore be stressful and tense, particularly if program staff do not

comprehend the nature or need for empirical research. Unforunately, this is too often the case (Nunnally and Wilson, 1975).

4. Feasibility

As discussed in the preceding phases of the evaluation, there may be several times when the feasibility of a good evaluation becomes questionable. Administrative constraints may prohibit developing a research design that truly enables the evaluator to answer the questions being asked. However, feasibility is not an all or nothing criterion. There are degrees of feasibility, and it is left to the evaluator to try to assess to what extent the research design is possible, and how it may be changed to accomodate the administrative problems which are evident.

5. Flexibility

Due to the possibility of the problems just discussed, it is important for the evaluator to be flexible and permit modification should the unexpected occur. In the real situation, there are a number of unforeseen problems which may arise and jeopardize the methodology and subsequent usability of results. It is therefore best to at least anticipate some problems, even thought unspecified. For although in the ideal situation, Murphy's Law is not prevelant, in the real situation, it always is.

c. Confidence Factors

The confidence factors of major import in any research are validity and reliability. Both are extremely important with regards to the research design and specific methodological procedures, which in the process model are identified as steps two and three respectively. Step three is just an extension of step two, and thus they should not be

considered mutually exclusive. However, they are distinguished from one another in the model, as there are different forms of validity and reliability which can be discussed relative to each.

One of the basic issues related to the choice of research design (e.g. experimental, quasi-experimental) is validity. Cook, Cook and Mark (1977) have identified four types of validity which should be considered in research design: Statistical conclusion validity, external validity, construct validity and internal validity.

Statistical conclusion validity deals with the validity of conclusions about the statistical association of a presumed cause and presumed effect. Conclusions are valid when assumptions about the statistical model are met and when the number of statistical comparisons made is adjusted for. Statistical conclusion validity can be increased by:

- having large sample sizes
- decreasing extraneous sources of error (homogeneous population and standardization of measurement)
- accounting for extraneous sources of variance (blocking or covariance analysis)
- increasing reliability of outcome measures
- standardizing implementation of treatment.

Statistics play an important role in social science research, and it is suggested that statistical inference is either overused or misused frequently (Beck and Brewer, 1978). With too heavy a dependence on statistical significance, what is meaningful becomes construed. As Deming (1975) suggests, statistical inference ends with the frame and environmental conditions under which the frame was studied. The theory

of probability cannot help outside these limits. The challenge related to statistical validity is not whether there is an error in the data, but whether the effects or error on the findings may be minimized, so that the meaningful variance can be given a chance to show itself (Eber, 1975). Multivariate techniques are the most sophisticated available for this process, for they capitalize upon chance less than repeated univariate analyses of the same data.

External validity is the extent to which a causal relationship can be generalized to or across persons, settings or times. Choosing heterogeneous groups of persons, settings or times therefore increases external validity. Bernstein et al (1976) suggest that there are a number of situational variables, such as the effect of the staff, Hawthorne effects, novelty effects, history, the geographical setting, etc. which can have a profound effect on external validity. Unfortunately, too little emphasis has been placed on external validity in evaluation research (Bernstein, et al, 1976).

Construct validity relates to inferences which are made about constructs on the basis of particular manipulations and measures. It is important to demonstrate that different measures of the same presumed construct covary and measures of related but different constructs to not. Construct validity is crucial in evaluation research (Tabor, 1978) for unless the indicators measure what they are supposed to, namely the concepts derived from the program objectives, evaluating the success of the program may be based on irrelevant criteria, or relevant criteria which remain unidentified and therefore not understood. The strength of a treatment is also an issue in construct validity, for it relates to the size of an effect related to specific variables. Moreover, it is

important to consider negative side effects of the treatment, in weighing treatment effects in terms of costs and benefits. Erroneous conclusions are also possible, such as an effect is assumed to have occurred, when in fact it has not, or an effect is assumed to have happened for the wrong reasons.

Internal validity is the validity of conclusions about whether the statistical association of a treatment as implemented and an effect as measured can reasonably be considered as a causal association. This type of validity has received the greatest amount of attention and discussion in evaluation literature. In fact, Sechrest (1979) contends that internal validity may have received a disproportionate share of attention to the detriment of more basic construct validity of treatments. He suggests that the planned strength of treatment and integrity of treatment have important implications for construct validity.

The confidence factor of reliability is one of the degree of variation in the measured phenomenon due to inconsistencies in measurement, rather than in the phenomenon being measured. Factors which may influence random error include:

- characteristics of the respondents
- characteristics of the researcher
- conditions under which measurement is made
- problems of measurement instrumentation
- problems of data processing and coding (Rutman, 1977).

Since evaluators of program effectiveness draw conclusions about programs and their effects, there is a great concern about assuring the highest degree of reliability and validity of the measures (Rutman,

1977).

Considerations of these factors before and during the research design and program specification can prevent later problems and mitigate the effects of disappointing and/or questionable results. In addition, attempts to explain these factors to the program planners can possibly lead to a greater understanding and appreciation of research, by the program planner. If, as expected in the ideal situation, the planners and evaluators are cooperatively developing the most useful strategy for evaluation research, consideration of these factors is critical.

2. Research Program Design

Development of evaluation research design is one of the more controversial areas of discussion, particularly with respect to the pros and cons of the experimental paradigm. Bernstein et al (1976) report that in a review of methodological procedures used by federally funded evaluation research in fiscal year 1970, research is lacking in the application of appropriate design, sampling procedures and data analysis techniques. Much of the debate focuses on the selection of experimental, quasi-experimental and other types of research designs. Other types of research include observation, participant ratings, or pseudo experiments such as the one-treatment, one-group, post measurement design (one of the least preferred (Nunnally and Wilson, 1975)). Pseudo experiments require multiple measurements and tests for convergent validity if any confidence can be established in the conclusions (Edwards, Gutentag, and Snapper, 1975).

The major difference between experimental and quasi-experimental is randomization of subject assignment to a treatment and control group, thereby eliminating alternative explanations for any change or effect

(e.g. internal validity). In the quasi-experimental design, some potentially important confounding variables are not controlled. Considerations for randomization include (Connor, 1978):

- Who will control the random assignment?
- Which is better: fixed random assignment or variable random?
- Is it preferable to group or block clients before random assignment?
- Should the researcher inform clients of the procedure?
- Will there be compensation for the control group?
- Who and how will the assignment be monitored?

The trade-off is between threats to internal validity and feasibility and practicality. In many situations, a true experimental design for research is either impossible or impractical (Alwin and Sullivan, 1976). The evaluator must therefore rely on naturally occurring treatment and control groups ex-post-factor. Boruch and Rindskopf (1977) however, suggest that randomized tests are more feasible in the social sector than one might expect, and efforts to capitalize on opportunities for experimental designs should be taken advantage of. Weiss (1972) has reported that the experimental model has come under attack not only because it is not feasible, but because it is counterproductive. It is necessary to hold the program constant, rather than attempting to facilitate improvement in the process. Too many controls and too many conditions can make the program ungeneralizable to the "real world", evidence of the trade-offs between internal and external validity. Weiss (1972) suggests that quasi-experiments have the overriding virtue of feasibility and their use has been considerably more frequent in evaluation research during the past ten years.

The issues of concern that differentiate these methods are those of methodological rigor which enhance internal validity and provide confidence in conclusions drawn from the data, and those of feasibility and generalizability. The decision is of course, based on a number of factors, and so, a combination of research designs may result.

3. Methodological Procedures

Special attention must also be given to the development of methods of sampling, measurement and data analysis within the research design. Too often, research is conducted with less than adequate sampling, design and measurement (Cook and Gruder, 1979). A good outcome measure is one which is feasible to measure (Rossi, Freeman and Wright, 1979) and thus, constraints of time and budget are again relevant. Accurate, valid, and reliable measurements of outcomes are also critical. Unreliability in measures can dilute and obscure any real differences when they do exist. Validity is also an important evaluative criterion of measurements, particularly with regards to predictive validity, content validity and construct validity.

With regard to data analysis, it is important to contemplate the procedures for analysis prior to design decisions, to insure that there are adequate statistics available to provide a meaningful analysis. In addition, as Eber (1975) has pointed out, a researcher must analyze data from two standpoints: his own in an attempt to achieve complete and rigorous understanding of the results; and from the program planner perspective, in order to better communicate the relevant parts of those results to the consumer. These two approaches may well require complimentary but different statistical models.

B. THE REAL

As previously discussed in the Phase II goal specification phase, Chicago Spacewatch was already a design and goal specified program when the evaluation was proposed to NSI. Thus, the important factors of influence related to the goals are very weak, ambiguous, and without due consideration in the research design. Without a clear understanding of the purpose and goals of the program, I experienced great difficulty in developing a research program of value to NSI. NSI was looking for data to support the premise that Chicago Spacewatch was a productive program, worthy of more funding from large aerospace industries, and the type of program which should be implemented elsewhere in the U.S. I erroneously concluded that they really wanted some good information on how to improve the program for future use. Whereas NSI was looking for an impact assessment to justify requests for additional funds, I was intending to look into the process of Chicago Spacewatch to greater understand any impact which did occur. Therefore, the research design did not include any in-depth investigation of a single program, but rather a general analysis of the process through which the program operated to accomplish its goals. It was thus proposed, that the greatest potential of the evaluation would be in its future applications. Unfortunately, due to inadequate attention to Phases I and II of program evaluation previously discussed, the program evaluation was working in cross purposes with NSI's intentions to conduct a program evaluation. NSI's purpose was to get more money for more programs. My purpose was to identify problems, and evaluate the process in order to provide data useful in developing better programs. The research plan was developed on that basis.

1. The Research Plan

During a two week period, I prepared an evaluation proposal on the basis of the predetermined goals and a brief description of Chicago Spacewatch activities. My research plan consisted of three phases, each related to a different aspect of the program, different individual programs, and different goals. A summary of the program activities, research and results of each phase are provided in the Executive Summary of the Chicago Spacewatch Program EValuation (Appendix A). Discussed herein are my intentions and rationale for research activities for each phase as they were at the time of the proposal. There were some changes made to the proposal during program implementation which will be discussed in the Program Implementation Phase.

a. Phase I: The Overall Program

The overall program goal of Chicago Spacewatch was to increase awareness, interest and understanding of space and space technology applications. Therefore, the first measure which was proposed was a base-line measure of the general space knowledge of the general public. This included an assessment of such factors as NASA spinoffs, NASA budget, R & D activity, technology applications, etc. A sample of the Chicago public was to be asked to complete a questionnaire which would measure their awareness and knowledge of space activity and its applications.

The theoretical assumption was that by increasing awareness, interest and understanding of space and space technology applications, a more pro-space attitude would be evident. Since this was assumed to be a focus of the overall program, it was decided that attitudes of the space program, and any changes in attitudes which may result as a

consequence of the Chicago Spacewatch program, should be empirically measured to test the assumption. It was also proposed that a baseline attitude survey be conducted with the same sample of the population to determine to what extent knowledge of space activity is correlated with a space attitude. Although there had been insufficient time for a full literature review regarding this topic, there were a couple of NASA studies which eluded to the fact that this may be the case (LaPorte and Metlay, 1975). Since this assumption was central to the focus of the program, this measure was considered to be important to the evaluation.

To determine whether Chicago Spacewatch Community Programs (e.g. magazine articles, TV programs, radio debates, museum attendance) had any influence, a community survey of space attitudes and knowledge was proposed with a sample of 1,000 Chicago residents, measured before Chicago Spacewatch and 1,000 measured after the program. The same instrument would be used in the post test phase; however, the post program questionnaire would include questions regarding the individual's awareness of and participation in Chicago Spacewatch activities. It was proposed that the results of Phase I would provide the following:

- baseline information of general space awareness and attitude.
- information regarding the change in public awareness and attitude possibly resulting from the efforts of Chicago Spacewatch (subject to interpretation of all measures together).
- correlation of space knowledge and space attitude (e.g. what particular kinds of knowledge are correlated with a particular attitude).
- information regarding advertising effectiveness.
- information regarding participation in public programs and its

effect on space knowledge and attitude.

b. Phase II: Community Programs

Phase II of the proposed research design addressed the community programs of Chicago Spacewatch. The intention in this phase of research was to evaluate public participation in specific community programs in order to quantify the extent to which Chicago Spacewatch would reach the people. It is first necessary to get potential program participants into the arena of activity before the program could be implemented and have its effect. The primary function of this phase of research was to determine whether or not advertising was effective in reaching the people, to determine if they heard of Chicago Spacewatch and attended any of the programs. The following research activities were proposed:

- a periodic telephone survey conducted with a random sample of the Chicago area population. A short, structured interview would be used to request information regarding special TV program viewing, attendance to the Museum and Planatarium exhibits, celebrity appearances, etc.

- wherever available, demographic and archival data providing characteristics of the population reached by each of the individual programs would be collected (for example, the number of people attending the Museum during Chicago Spacewatch as compared to similar dates in previous years) to provide some indicators as the effectiveness of advertising.

- two community programs focused on the elderly population (the reason for this focus was that Hugh Downs, the President of NSI, was developing a series on the elderly for his show 20/20. It was felt that

some data should be collected from this unique group and thus a short interview was designed with the intent to interview participants at these two programs. My major objective was to obtain feedback from the elderly participants regarding their perceptions of the overall program (e.g. did they enjoy it, what did they like best, did they learn anything, etc.). There was no intention to measure a change in attitude toward the space program, for the programs were too diverse in nature to provide any direct focus on the space program.

- interviews with speakers and administrative of the individual programs were, to obtain their reactions to and impressions of Chicago Spacewatch, as well as opinions of program proceedings. The results of Phase II were intended to provide:

- record of attendance of all public activities
- comparison of special exhibits attendance in relation to other similar points in history
- subjective evaluation by participants of the Elderly Persons Luncheon (one of the elderly programs).
- random survey of public activity attendance
- information regarding advertising effectiveness.

c. Phase III: Educational Programs

In Phase III of the evaluation, the research and data collection were focused on the educational programs, specifically the teacher programs. The teacher population was considered primary in this aspect of the research. Three groups of teachers were the target subject populations:

1. teachers attending the teacher orientation at the Adler Planetarium.

2. teachers invited to attend the orientation and who did not attend.

3. teachers in similar school systems not invited to attend.

All three groups would be asked questions relevant to the assessment of a space attitude and space awareness, by means of the same instrument used for the general public. It would then be possible to compare attitudes and knowledge of teachers with the public groups. It would also be possible to compare the attitudes and knowledge among each of the three groups of teachers. I hypothesized that if a teacher attended the orientation, it would be because he/she is more interested, and thus would demonstrate a more positive attitude and possibly more knowledge of the space program. Group II would be asked why they did not attend the orientation, and Group III would be asked if they would attend such an orientation if invited.

Pre and post measures were proposed for Group I teachers. At the beginning of the orientation, a questionnaire would be distributed for completion and those same teachers would be sent a second questionnaire nearly one month following the orientation which would be designed to obtain in addition to a second measure of space attitude and awareness, information regarding utilization of information obtained from the individual programs. In addition, a sample of teachers who would have speakers attending their classes, would be sent a questionnaire by mail, asking for ratings of their speaker and presentation. The results of Phase III of the evaluation were therefore intended to provide the following:

- information of program attendance and utilization.
- information of change in teacher attitude and awareness as a

result of their participation in Chicago Spacewatch.

- additional correlational knowledge of space knowledge and attitude.
- sample ratings of speaker programs.

These were the research activities suggested to NSI in the proposed evaluation. NSI was informed that there was no intention to provide an in-depth investigation of any single program, but rather a general analysis of the process through which the program would operate to accomplish its goals. It was stressed, therefore, that the greatest potential would be in its future applications. In summary, the proposed evaluation research design was intended to provide the following:

- indicators of goal achievements and program effectiveness
- input for future decision making
- specifics for future funding considerations
- instruments tested for reliability, reuseable in future programs
- baseline data of space awareness and space attitude for use in future city comparisons, as well as in a comparative analysis of program effectiveness
- a unique stimulant, in that the process of evaluation in itself, is an awareness generating technique.

Whether any or all of these intentions were to be realized was not known at that time. The overall program activities were intended to provide the baseline data of space awareness and space attitude, which was thought to be of considerable important in future decision making. It could provide a standard from which to build information dissemination programs. The evaluation activities for the community and

educational programs were intended to provide more specific information, such as the indicators of goal achievement and specifics for future funding considerations.

2. Discussion of the Research Plan

As previously discussed in the ideal section, the first step in the process of developing a research program is the identification and consideration of several factors of influence, all of which are relevant in any research design, but seem particularly so in mine. With regard to goal related factors of influence, it should be evident from previous discussions that there was very little input from such factors as the goals of the programs, the questions which the planners wanted answered, or the criteria for success. This type of information simply was not available.

Time was a very critical and crucial element in the development of the Chicago Spacewatch evaluation. The proposal was put together in one week's time on the basis of only a bare skeleton of the proposed program activities. The next week a decision was made, and the next week it was absolutely essential to start collecting pre-test data prior to the implementation of the program. The time factor is important for several reasons:

- the research design was not well thought out. There was no time to confer with consultants or review possibilities with my advisors.

- there were constraints in the consideration of using random samples. Telephone interviews with a random sample of the Chicago population would have been preferable to surveys ompleted on the streets. However, the latter was accomplished in two days, whereas the former would have taken approximately two weeks.

Another constraining administrative factor of influence was the budget. Nearly every dollar was spent in the manner intended, and I was careful not to exceed allotted funds. However, it is important to realize that I provided an estimated $8,000 evaluation for the cost of$1,800, covering expenses only. I had been given a figure to work with, and developed a research design in accordance. Thus, both time and money were major constraints.

There is another relevant factor of influence that may possibly be categorized as administrative. It is in regards to perceptions of my "role" as an evaluator. I experienced difficulty in maintaining two roles simultaneously: that of the student and that of the professional. NSI was completely aware that the evaluation was to serve as a master's thesis, and thus, it was evident that I had not obtained a PhD. This can lead the administrative staff to conclude that they know more about what you are doing than you do. Although, at times that may be true, the authority structure in evaluation decision making was often unclear. It become difficult to make decisions and pointedly ask questions which may be embarrassing to the administrators. Defensiveness coupled with authoritarianism was potentially disastrous. This is not to suggest that NSI personnel were uncooperative. In fact, the opposite was true. There were times however, when I felt that they were not taking my work as seriously as I thought they should have.

The confidence factors necessary for consideration in research design (validity and reliability) were definitely considered; however, I must admit they were not considered and acknowledged to the extent that they should have been. Internal validity was the major problem. My design consisted primarily of pre and post measures, which are subject

to a number of threats to internal validity, such as history, maturation, testing, selection and mortality. In the educational programs, I attempted to establish some control by including a comparison group of teachers who had not participated in any activities. This was helpful in drawing conclusions. My sampling procedures were also questionable and could have presented problems in the interpretation of results.

There were, however, such significant problems in program implementation, that my research design did not become of critical importance in the interpretation of results, as will be discussed later in the Program Implementation Phase. Had it been, it is possible that the results of my research would have been subject to a number of serious threats to validity. Interpretation would have been considerably more difficult on my part, and more easily misunderstood on the part of NSI.

In step 3 of the ideal process model, measuring instruments and data analysis are also important considerations. The primary instrument I used for measuring space awareness and attitude was a questionnaire, developed through a thorough process of pilot testing. The procedures in the pilot test included a telephone survey of 25 Chicago residents, who provided information necessary to compose questions. Several items were generated which either asked a question about the space program or required an attitudinal response about the space program. A pilot questionnaire was tested on 85 Loyola Psychology students. The responses were factor analyzed to determine which items had the highest reliability with the overall test and to determine which items collectively identified a particular dimension or factor. Those items

which produced the highest reliability and discrimination in the first pilot questionnaire were incorporated into the second pilot questionnaire administered to 50 Loyola undergraduates and 30 Chicago people in the downtown area. The resulting basic questionnaire used in the evaluation consisted of 21 Likert items for measuring a space attitude, six questions measuring knowledge of space related activity, and eight questions related to whether respondents thought specific benefits were a product of the space program.

Data analysis was carefully considered at the onset. Plans were made to obtain consultation with statataicians familiar with multivariate analysis, which was considered to be the correct means of evaluating Chicago Spacewatch data. Statistical validity is not considered to be a problem in my design.

In many cases, the specifics of the design proceeded and became formalized as the program moved along. For example, although the questionnaire used for the teacher orientation was developed early, later questionnaires which were sent by mail to other teachers were developed at the time in which they were needed. Efforts were made to incorporate in these later questionnaires Fishbeinian theory, including questions related to subjective norms, behavior intentions, etc.

In summary, the Chicago Spacewatch evaluation research design was adequate to meet the needs of the evaluation of the program as it was implemented; however, in comparison to the ideal process of formulating a research design, the Chicago Spacewatch evaluation research design was replete with problems. With regard to factors of influence, goal factors were not addressed, administrative factors were considered and found to be rather constraining, and confidence factors related to

internal and external validity were not high. The choice of designs (i.e. experimental, quasi-experimental, other) was made too quickly and without sufficient consideration of alternatives. Sampling procedures for the general population survey data were also a problem potentially, in that if any differences had been found between pre and post measures, selection may have become an important factor for consideration. However, since no differences were found, the data was accumulated across time to provide a general understanding of the correlations among space awareness, space knowledge, and space attitude.

Although it is probably true in every case that with more time and planning, a better evaluation can be developed, this is particularly true with the evaluation of Chicago Spacewatch. Most crucial, is the time necessary to understand the planner's purpose and goals of the program to ensure that the evaluation provides the most relevant information to the planner.

CHAPTER VII

PHASE IV: PROGRAM IMPLEMENTATION

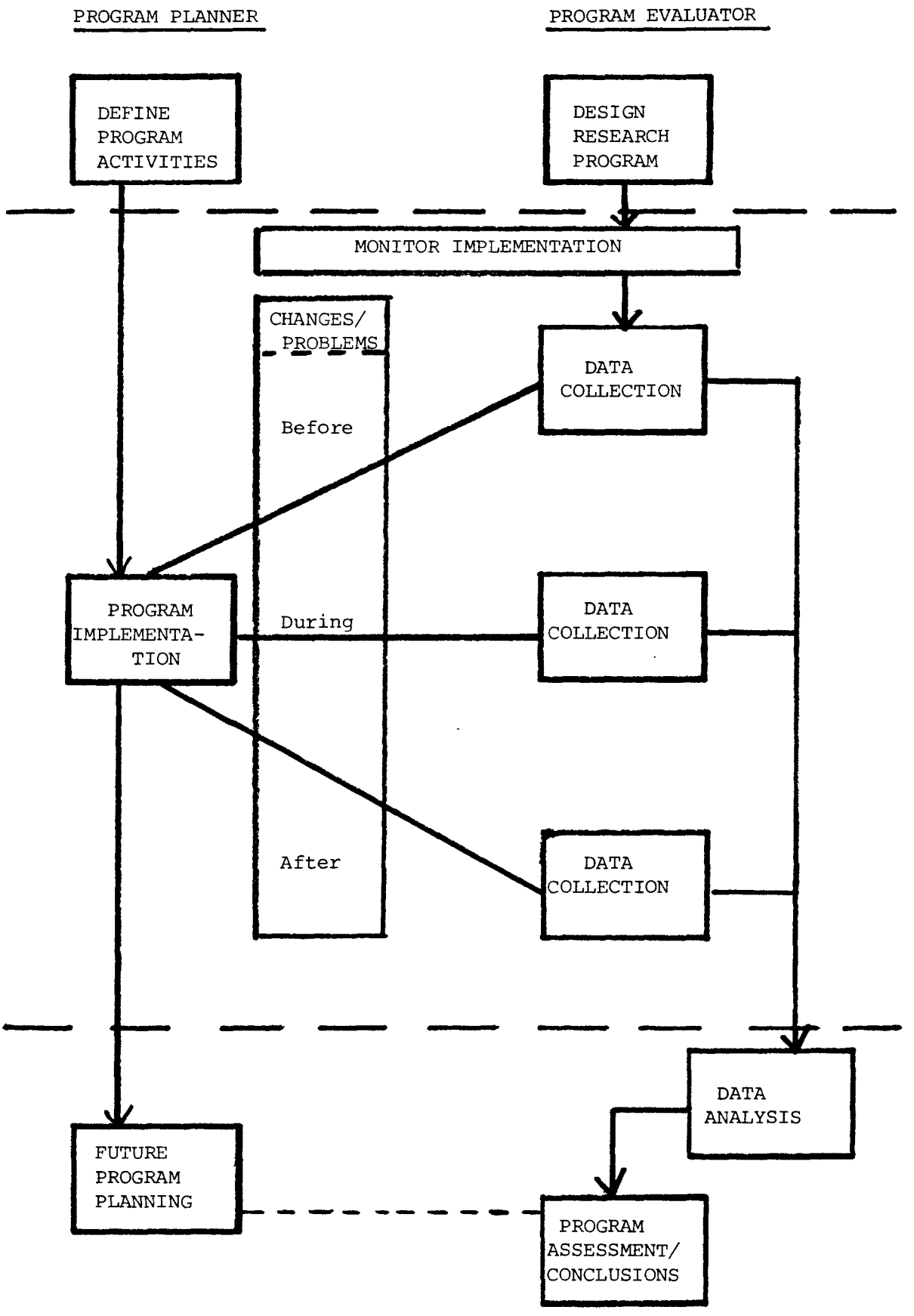
A. THE IDEAL

The actual implementation of the program(s) can become an evaluator's nightmare, if unprepared for possible program changes. Cognizance of the famous Murphy's Law is essential. Sechrest et al (1979) suggest that when programs are very complex, requiring the delivery of a wide variety of services from diverse agencies, apprehension should be strong. Boruch and Gomez (1979) state: that a new social program will be implemented imperfectly, is obvious once said.

Evaluation methodology often limits attention to those outcomes which fall under the stated goals (Rutman, 1977); limiting the scope of the research, and possibly missing latent goals, unintended consequences, and other unanticipated effects. Therefore, a major purpose for monitoring the program's operation is to determine whether there is uniformity in the implementation of activities, as prescribed prior to program implementation. Figure 5 is the ideal process model of this phase of the evaluation, and as indicated, the major question is whether the program is implemented in the manner in which it was prescribed prior to implementation. To neglect this area according to Patton (1979) is to create a black box between the periods before and after treatment (e.g. program implementation), the secrets of which are critical to an understanding of the results. Unless the evaluation can provide evidence on the nature of the program as it existed, there is

FIGURE 5

PROCESS MODEL OF PHASE IV: PROGRAM IMPLEMENTATION



little basis for a decision (Weiss, 1972).

As indicated in the process model, there are changes which can occur before, during, or after program implementation. Not only are there many potential changes which can occur in the implementation, but a multitude of problems may arise as well (Murphy's Law). If rigid experimental controls are necessary for the research design, it should be evident, that changes/problems affecting the program are not welcomed by the evaluator. Edwards, Guttentag, and Snapper (1975) suggest that there are four kinds of changes which occur as the program progresses:

1) Values of both those served by the program and the program people change in response to experience with the program and in response to external causes.

2) As the program evolves, it will change in shape and character.

3) External societal circumstances, to which the program is a response, will change.

4) As knowledge of program events and consequences accumulates, changes are made in response to this knowledge.

In addition, programs may be changed and/or influenced by budget cutting or budget expansion, changes in administration officials, veering of the ideological winds, change in congressional support, public appraisal, initiation of rival agencies and rival programs, pervasive client dissatisfaction and critical media coverage (Weiss, 1975). There may also be failures in the delivery of services, in that no treatment is delivered at all, the wrong treatment is delivered, and/or the treatment is not standardized, is uncontrolled, or varies across target populations. And there may be the problem of delivering negating treatment, problem of creasing (when treatments are given to

those most likely to benefit), problem of ritual compliance (treatment is watered down), problem of overly sophisticated treatments, problem of client heterogeneity, and the problem of client rejection of treatment (Rossi, 1979).

With just this meager list of potential problems and changes, it is evident that the evaluator must be aware of all of these things if appropriate interpretation and accurate judgements are to be made regarding program effectiveness. If the program is vague or novel, or being developed as it goes along, the evaluator may need to describe what is going on to clarify the meaning of the program, and to contribute analysis of which features of the program work and which do not (Weiss, 1972). In addition, unless the evaluation can provide evidence on the nature of the program as it existed, there is little basis for a decision of effectiveness. Problems and changes in program implementation are particularly disconcerting to an evaluator if an experimental design is used, for the researcher must somehow control implementation without allowing the research component to unduly affect implementation and outcome (Twain, 1975). In this context implementation control may be difficult, if not impossible.

According to Patton (1978), evaluation research has been dominated by outcomes assessments, with considerable attention placed on goals clarification. However, if one had to choose between implementation information and outcomes information because of limited resources, there are many instances in which implementation information would be of greater value; for one obvious reason: until the treatment is implemented, there is little reason to evaluate outcomes. The study of implementation ideally should not supplant evaluation of outcomes;

rather the two research efforts should be complimentary (Weiss, 1972).

Patton (1972) and Williams (1978) suggest an approach to evaluation with implementation as the focus. Williams (1978) suggests implementation as a research question, in which an inquiry about implementation seeks to determine whether an organization can bring together people and material in a cohesive organizational unit and motivate them in such a way as to carry out the organization's stated objective. Thus, an implementation analysis could investigate the technical capacity to implement, the political feasibility, and technical and political strategies for implementation. Unfortunately, it is rare for implementation to be considered or analyzed during the decision making stages of program development. Implementation success however, is a viable research question; for at some point there should be a determination of the degree to which an innovation has been implemented successfully. In addition, implementation success can be used as a covariate in impact analysis if a comparison group is available.

According to Patton (1979), it is important to frame evaluation questions in the context of program implementation. Patton argues that evaluation has been dominated by an emphasis on measuring outcomes and there are so many problems with outcomes that the results of an evaluation give decisionmakers very little information upon which to act. What is missing is information about the actual nature of the program being evaluated. What's happening inside the black box, the program? According to Patton, black box evaluations that study outcomes alone do so because of tradition and routine. Failure at the implementation stage may be a major reason for a shortfall in human

services and ineffective programs. A very serious challenge in program evaluation is the determination of how far from the ideal plan a program can deviate, and in what ways it can deviate, while still meeting fundamental criteria. In other words, how different can an actual program be from its planned ideal, and still be said to have been implemented? Patton (1979) suggests three types of implementation evaluation:

1) Effort evaluation: this type documents the quantity and quality of activity that took place. It is an assessment of input regardless of output.

2) Process evaluation: this type focuses on internal dynamics and actual operations of a program in an attempt to understand its strengths and weaknesses. Relevant questions would be: Why are certain things happening? How are the parts of the program fitting together and how are people perceiving the program? It requires a sensitivity to qualitative and quantitative changes in the program, from a variety of perspectives including the evaluator's perspective, and perceptions of people close to or involved in the program.

3) Treatment specification: this type of evaluation would identify and measure precisely what it is about a program that is supposed to have an effect. One would measure the degree to which specified treatments actually occur, which can be difficult. However, it can reveal causal assumptions underlining program activities.

That a program be implemented is obviously critical, for if conclusions are going to be made regarding the effect of a program, there must first be a program. Moreover, an understanding of the potential changes and problems which can occur during program

implementation is a valuable asset to the evaluator, and flexibility can become a meaningful and desirable trait. Unfortunately, however, it has been known to happen that a program has not been implemented and yet evaluators have proceeded with their research without careful monitoring of the implementation.

The discussion of the ideal situation during this implementation phase has focused on the program planners, and their changes and problems associated with the implementation of the program. However, it should be noted, that for all said with regard to the program, the same is true with regard to the actual implementation of data collection procedures in the evaluation. The evaluator assumes a responsibility for a different type of program, the research program, which although independent of the program being evaluated, is susceptible to similar types of problems - budget changes, insufficient help, too little time, etc. In addition, the research program must be responsive to and reactive to changes in the program being evaluated. In the ideal situation, evaluators and planners are working together and thus, the possibility of an evaluator evaluating a program that is not implemented should not be possible. However, in reality, it is.

B. THE REAL

In the beginning, I had no doubts that the programs of Chicago Spacewatch would be implemented without any significant problems. In fact, I had great expectations of NSI hitting Chicago with a storm of influence and intrigue, stirring the population with excitement and curiosity in regards to the future and the space program. I had expected announcements of Chicago Spacewatch in all newscasts, front page articles in the newspapers, astronauts visiting throughout the

town, elderly lips smacking with astronaut food, and children sparkling with new enthusiasm. I had expected extensive media coverage and the Chicago population smothered with Chicago Spacewatch programs. I had expected too much and with great expectations, reality strikes hard.

As related to the process model previously discussed (figure 5), the changes and problems which became most apparent and most influential in the potential significance of my research program occurred "during" the program. Measures which I had intended to take before the program (1,000 community surveys and a pre-test of teachers attending the teacher orientation) were done prior to the end of September. On October 1st the ribbon cutting ceremony was conducted with community pre-tests finished. Chicago Spacewatch was underway, and shortly thereafter I became aware of some serious problems in the implementation of Chicago spacewatch.

The most major problem relates to Chicago Spacewatch as a program entity, consisting of a series of educational and community programs coordinated by the National Space Institute. During the program implementation, I learned that the idea of a "Chicago Spacewatch" existed long before NSI had any role in its development. Chicago State University (CSU) had previously coordinated a space program activity with NASA, to be centered around NASA's 20th anniversary. CSU had developed a number of programs for the educational community, calling their program "Chicago Meets Outer Space" (CMOS). The major focus of this program was on minority groups and the role they might play in space program development. Throughout Chicago Spacewatch, CMOS maintained its own identify in all programs held at CSU. The confusing factor to me, as the evaluator, was that I had assumed that

NSI was the coordinator of all programs, responsible for the assemblage of all these individual programs into the overall program, Chicago Spacewatch. I further assumed NSI had control over the proceedings of each individual program. These assumptions were incorrect. Although NSI was attempting to place a coordinating cover on all activities, most of the activities maintained their independent identities and found it extremely difficult to coordinate their plans within the overall schema imposed by NSI.

For example, CMOS had a week long program agenda developed including:

- a teacher workshop where teachers could earn extra credit during a one week evening class with NASA instructors. The focus was on how to teach space in the classroom.

- community activities on the weekend, with special programs on astronaut food, NASA exhibits, computer demonstrations, etc.

- educational programs during the week for the students from the City of Chicago. Astronauts visited and NASA provided talks and lectures.

- special ribbon cutting ceremony with attendance by political figures.

Many of the Chicago Spacewatch activities overlapped with those of CMOS, causing confusion and ultimately poor attendance at many of the individual programs. Interviews held with personnel at CSU confirmed my perceptions of displeasure with the role of NSI in the overall proceedings. For example, NSI had assumed responsibility for press coverage of CMOS events, however, only those activities that were independently promoted by CSU were well attended.

The teacher programs are a good example of the program redundancy and lack of coordination of activities. Teachers with the Chicago Public Schools received three separate announcements (independent of one another) announcing the teacher workshop, the teacher orientation, and the special program for visiting speakers who would attend the classroom and discuss space topics. Some teachers informed me that they found this very confusing; others expressed annoyance (via the questionnaires). It is probable that the overlapping of both programs and advertisements is a contributing factor to poor attendance to the teacher workshops. Several teachers also made contact with me as a result of my questionnaires, asking me to straigthen out a mess in which no one seemed to know who was in charge of the speaker agenda. I didn't know either. As the evaluator, I was simply unaware of the "real" circumstances surrounding the development of the program "Chicago Spacewatch".

Another example of independent functions was the O'Neill presentation held at Navy Pier, in conjunction with a student debate on space exploration. The Chicago Society for Space Settlement (CSSS) independently sponsored this activity. This organization paid the speaker, prepared and distributed all announcements and promotional literatures, etc. The attendance exceeded 300 (considered very good by CSSS), yet very few attendees were aware of the program "Chicago Spacewatch". This may be partly due to the fact that "Spacewatch" is the name of the CSSS monthly newsletter, the origin of the Chicago Spacewatch name.

These few examples demonstrate that NSI was not a controlling force in the planning and coordination of activities of Chicago

Spacewatch. In fact, the program Chicago Spacewatch was really never implemented; only a series of educational and community programs independently and within close temporal and proxemic contiguity were implemented. At no time was it clear to the majority of attendees at any individual program that this was one of many programs within the overall Chicago Spacewatch program. Only on rare occasions did the words "Chicago Spacewatch" even appear in advertisements or announcements. Thus, although these independent programs did exist, they were not tied together collectively, which was the responsibility of NSI.

There is difficulty in evaluating a program called Chicago Spacewatch when it is never collectively referred to as such. Each independent participating organization had a different idea of what Chicago Spacewatch was. To members of CSSS, Chicago Spacewatch consisted of the student debates and O'Neill presentation. To CSU students, Chicago Spacewatch was the Chicago Meets Outer Space Program. Chicago Spacewatch did not exist. It was an abstraction, conceived by NSI to coordinate a series of activities, an action that was not satisfactorally implemented. Attempting to evaluate an abstraction is very disconcerting, and during the implementation of Chicago Spacewatch, I began to realize I had a very serious problem. My research design was intended to measure the effects of the overall program, not each individual program.

Implementation was also a problem with regards to several individual programs. The worst example and one which clearly illustrates the problem was the special community event, Senior Citizen's Day in Chicago. This was one of the first activities of

Chicago Spacewatch that I was to attend. Senior Citizen's Day was a special event, and NSI was to play a minor role, presenting a lecture of the benefits of space to the elderly. Exhibits were to be set up with space food, demonstrating the adaptation of space food into freeze-dried nutritious food for the elderly. NSI's role in Senior Citizen's Day consisted of NASA exhibits at the front entrance of St. James Cathedral (where it was held), a speech given by Tom Gorski on the benefits to the elderly from the space program, and a display of astronaut food adapted for the elderly.

On the basis of my observations and interviews conducted with the elderly participants, I would conclude that this program had major problems in terms of implementation. As the elderly entered the front door, none of them stopped to see the exhibits, paid no attention to them, and perhaps did not even see them, as they were set over to one side. The exhibits were not noticed by a single individual in a half hour's time during which observations were made. Gorski's speech was delivered during lunch breaks and very few elderly heard it or for that matter paid much attention to it. Their focus of attention was on the singing and dancing and other activities which preceded and followed the Gorski presentation. I observed considerable inattentiveness, as did Gorski, and also as did others in charge. Regarding the astronaut food display, it was not adequately explained to the elderly exactly what they were eating when they were offered a sample taste. I interviewed many elderly, asking them what they thought of the astronaut food. Most of them thought it was their lunch and they were concerned about the small "stingy" samples. Interviews with the elderly about the space program in general, provided some meaningful data with respect to their

general opinions about the space program. However, the role of NSI in Senior Citizen's Day was negligible. Most elderly, when I questioned them about the space program, with puzzled expressions on their faces wondered why I would be asking about the space program at such an event. Although it could be said that this individual program was implemented, it would be very difficult to say that anyone was aware of its implementation.

In many respects, implementation was a problem with several programs in Chicago Spacewatch. According to NSI, over 100 magazine and newspaper articles appeared in the paper during the month of October, related to Chicago Spacewatch. However, most of these articles were tiny little quips of insignificant detail and were never tied to the overall program. There may have been space articles, but the advertising and promotion which I had expected were simply non-existent. According to NSI they had difficulty in establishing liasons with the press and media in Chicago. Press relations should have been developed long before the actual implementation of the program, as that part of the program was crucial to its implementation.

It was clear after the first week of the program, that there were going to be some problems in establishing that Chicago Spacewatch had in fact been a program implemented. To confirm this suspicion, the telephone interviews conducted with a random sample of Chicago residents suggested that the majority (vast majority) of the people had not heard of Chicago Spacewatch, nor were they aware of the many programs which were being held throughout the city, such as special programs at the Museum or Adler Planetarium.

Questionnaires were also mailed to local members of NSI and two

other space advocate groups, the World Future Society and the Chicago Society for Space Settlement. These questionnaires asked respondents if they had heard of Chicago Spacewatch, what programs (if any) they had attended, and also asked about interests and activities related to the space program. Members of these groups were mailed questionnaires for several reasons:

- Members of the group were assumed to have a pro-space attitude. Members would therefore be the people who would most likely attend to information about the space program. Thus, by asking them if they had heard of Chicago Spacewatch, and what programs they knew of and/or attended, some evidence of advertising effectiveness could be obtained.

- As a member of a pro-space organization, it was reasoned that members would have a very positive attitude toward the space program, would know more than the average person about the space program, and would perceive more benefits from the space program. They would therefore provide a good check on the validity of the instrument to measure these variables.

- Other information related to their space oriented activities was requested as these data could be of use in requesting help from space organization members to assist in Chicago Spacewatch type programs elsewhere in the future.

As a result of these circumstances, additional tasks were introduced into the research design. The all purpose questionnaire, measuring space awareness, knowledge and attitude, was used in several settings in addition to the agreed upon four settings for pre and post measures. One of these settings was the Museum of Science and Industry during the weekend of several special Chicago Spacewatch activities

(e.g. John Denver visiting to announce support for the space program; a special Rockwell multimedia presentation; the elderly luncheon; Hugh Downs present for a news conference; R2D2 and its wonders). The questionnaire was completed by respondents entering the museum, and respondents who were observed by the surveyors attending any one of the special space features in the museum (e.g. shuttle exhibits, space movies, etc.). The questionnaire was also completed by the majority of attendees to the O'Neill lecture at Navy Pier.

At both the Museum and Navy Pier the questionnaire which was used included additional questions regarding respondent's awareness of Chicago Spacewatch. Did they hear of it? If so, how? The Museum questionnaire included such questions as: Why did you come to the Museum today? These surveys were intended to provide data which would help to evaluate the effectiveness of Chicago Spacewatch advertising and to determine the level of space knowledge and attitude of those people attending the programs.

Gorski was aware of my added efforts and supplied additional funds to cover the costs. He attributed the additional work to a hard working ethic. Unfortunately, we never had an opportunity to discuss the problems he and the NSI staff were experiencing with regard to implementation of the program. In addition to poor attendance at the programs and insufficient advertising, the press was not real responsive and commitments by astronauts and other celebrities to appear in Chicago during the program were broken. NSI itself had too little staff to monitor the implementation of all of the programs, and in many cases, they were dependent upon my reports for an appraisal for the program.

The major element omitted from the Chicago Spacewatch evaluation

(which according to the ideal process model is essential) was the intent to monitor the program implementation and be prepared for problems. As previously stated, my expectations were far too idealistic. However, I did prevail in the need to be flexible. Additional data were collected and I attended each program in order to understand more of the process of implementation to identify problems. The biggest problem was ineffective and insufficient advertising, and consequently poor attendance to the programs. As a further consequence, the evaluation research had a change in focus. The major focus at the onset was to determine the degree of change in attitudes, etc., as a consequence of the Chicago Spacewatch program. The focus was on the impact of the program. However, with the problems in implementation, acquisition of data which would reveal to NSI the sources of difficulties and problems became more important. The focus was now on the process of implementing the program, rather than the impact of the program. This change in focus was necessary in order to assist NSI. It should be noted, however, that all my actions were independent, without consultation or discussion with NSI staff. Although I perceived these problems and acted accordingly, it is not clear that NSI did. Again, lack of communication was a problem.

In summary, there were several lessons which I learned with respect to the program implementation phase of program evaluation.

- The evaluator should be realistic in his/her expectations. It is also important to be prepared for implementation problems.

- Both time and money should be planned into the research design and data collection procedures to monitor program implementation. It is a crucial part of the program. I began to monitor Chicago Spacewatch

programs only after the realization that not to do so would be detrimental to the evaluation. An evaluation of an unknown program can only lead to erroneous conclusions.

- Open communication channels between planner and evaluator are very important. In the ideal process model this is the case. In Chicago Spacewatch, this was not. There were very few opportunities for NSI staff and myself to actually sit down and discuss what was going on.

- It is critical to understand and be familiar with the intended process of implementation. This includes knowing who is in command, what the hierarchy of decision making is, etc. For example, in Chicago Spacewatch, due to the number of independent sponsors, I was at a loss to know who to consult until I had determined who the appropriate sponsor was of a given program.

- Related to this former point, it is advisable to identify, meet and interview periodically, all sponsors of the program. These interviews and contacts may help to ascertain who is in control of what functions.

- It is important to have a good comprehension of the program itself. It was most amazing to me to find out during implementation that Chicago Spacewatch was something other than what I had thought; not one large program, but rather a series of little programs, not tied together effectively. Had this been investigated before the program began, it is possible some changes may have been made or suggested to NSI regarding this approach.

In addition to the foregoing, the program implementation of Chicago Spacewatch made me most aware of the value of cooperative action between programmer and evaluator. Both can gain objectivity and

appreciation for another perspective from the other. Unfortunately, in the case of Chicago Spacewatch, the pressure of time obviated the opportunitites for close communication and contact. I feel, however, that had the ideal process model been adhered to in Chicago Spacewatch, both the program and the evaluation would have been substantially improved.

CHAPTER VIII

PHASE V: RESULTS/UTILIZATION

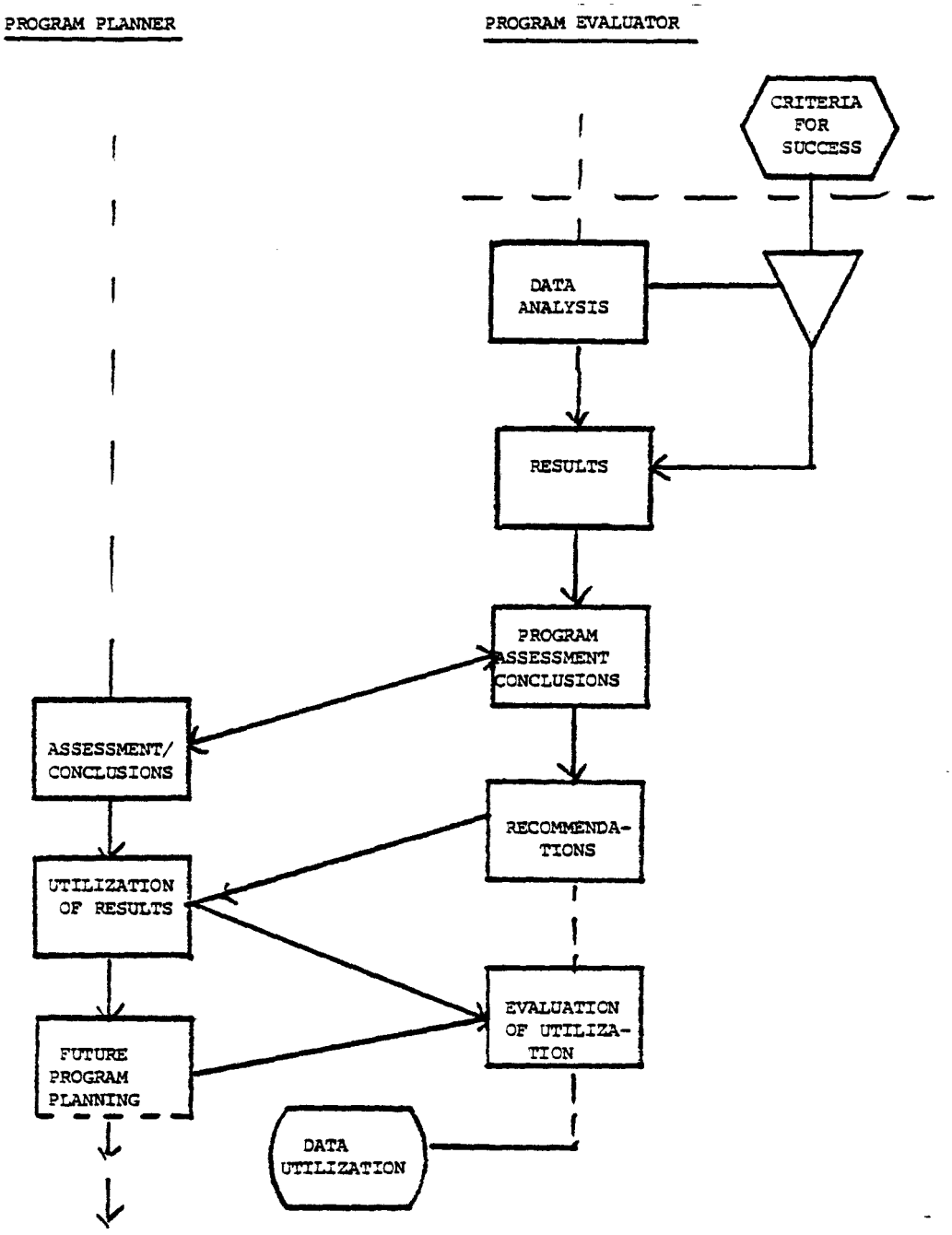
A. THE IDEAL

The final phase of the evaluation concerns the results of the evaluation and the utilization of these results by the program planners and to some extent by the evaluator. As indicated in Figure 6, after the data have been collected, the first step is data analysis. Data analysis is the reduction of data into some meaningful descriptive terms and statistical analysis which may provide the degree to which probability favors a cause effect relationship between treatment and outcome, or a correlation between variables. Data analysis includes organizing the data, constructing appropriate statistical tables, organizing data for display and reporting and making comparisons.

The results and program assessment/conclusions require interpretations, or making judgements about what the data mean, determining the implications of the findings, and linking evaluation results to future action. The results may be determined on the basis of a comparison between the resulting data and previously determined criteria of success, leading to assessments and conclusions. These functions are performed by the evaluator; however, in the ideal situation, both planner and evaluator should cooperatively make judgements and conclusions, encompassing both perspectives. It should be noted, however, that many decision makers and information users are not highly sophisticated about methodology and statistical analysis and interpretation, although they may be aware that methods and measurements

FIGURE VI

PROCESS MODEL OF PHASE V : RESULTS/UTILIZATION



are subject to question. Thus, the determination of conclusions involve trade-offs in which the validity of the data is matched against its relevance to the questions of the program planners. In the ideal situation, where planners and evaluators have worked together to design both programs and research, this aspect of the evaluation can be cooperative and exciting. In the real situation, however, it may be necessary for the evaluator to analyze the data from two perspectives - for the planner and for his/herself.

This is due to the different perceptions each may have of data analysis, different purposes for the evaluation, and in general, different backgrounds of each, as discussed in Section III. In addition, the evaluator is trained in data analysis and therefore is more cautious and cognizant of data misuse. The planner will then develop his/her own conclusions on the basis of the information provided by the evaluator. The evaluator may or may not make recommendations for future program planning. The use of all of the evaluation results is the major topic of discussion in this section.

The utilization of evaluation results has become a major topic for discussion and controversy in the field of program evaluation. Low utilization of evaluation results is a major concern to evaluators. However, there are a number of reasons why evaluation results may not be used, or may result in low utilization or underutilization. According to Agaruala-Rogers (1977), reasons for underutilization of evaluation research results are likely to include the following:

- lack of administrator's participation and involvement in the evaluation process

- conflicting interests of program officials and evaluators of the

program

- lack of mutually agreed upon "problem" definitions and "needs" deliniation between evaluators and users of the research findings

- lack of specialists who are trained to act as liaisons between the program administrators and evaluation researchers

- lack of emphasis on providing solutions to problems, other than accurate descriptions of events and activities only

- over emphasis on negative aspects of programs in the evaluation reports

- problems of feedback and timeliness of evaluation results.

The first five reasons reflect problems which have originated in phases and steps prior to this phase - during the research design or during the identification of purpose and establishment of goals. In the ideal situation, if these issues had been considered previously, they would not likely be influential in the utilization of results. However, in the real situation, these factors may not be worked out prior to the actual research. Over-emphasis on the "negative" aspects of the program is a real problem at this phase of the evaluation. There is a tendency for decision makers to use research only when its results match their preconceptions and assumptions in accord with their values (Weiss, 1975). Differences in values and value priorities constitute an inevitable limitation on the use of objective rational analysis. Program effectiveness, positive or negative, may be only one of the many values that enter the decision making process. The evaluation should be aware of these values; with good communication between evaluator and planner.

The last point of concern suggested by Agarwala-Rogers (1977) is

one of timing. Evaluation and program planning may operate on different calendars, resulting in the dissemination of results to program planners too late for profitable use. Evaluation reported a year or two or more after completion of the program is often too late to affect decisions. In some cases, data analysis time demands may constrain the evaluator; however, considerations of use may dictate at least some preliminary reporting of the direction of results in early phases.

Another reason for low utilization may be the politics of the situation. It is possible that evaluations may be disregarded if they address only official goals (Weiss, 1975). Evaluation must also assess the political goals and identify the measures most appropriate to measure political goals. Again, in the ideal situation, this circumstance would have been identified in previous phases and thus would not be a problem at this time.

There is also the potential problem of misuse of evaluation data such as in the following (Suchman, 1972):

- eye wash: an attempt to justify a weak or bad program by selecting data which "looks good" on the surface.

- white wash: an attempt to cover up program failure and/or errors by avoiding or preventing any objective appraisal.

- submarine: an attempt to destroy the program (torpedo), irrespective of its effectiveness.

- posture: an attempt to present the appearance of objectivity or professionalism without true regard for either.

- postponement: an attempt to delay action under the pretext of collecting data and seeking facts.

Mis-utilization may be hard to define and/or discover (Cook and

Pollard, 1977) and thus, may be subject to interpretation. Similarly the definition of utilization is subject to interpretation and is not easily defined. Cook and Pollard (1977) define utilization as a diffusion of results used to form a decision or justify a decision. However, it is important not to adopt too narrow a time perspective with this definition, in that it is difficult to establish the appropriate time lag between the finished results and utilization of the results (Cook, 1978). In summative evaluations, the results are more clear cut - the program either continues or not. However, in formative evaluations, changes may occur slowly and gradually address different aspects of evaluation results. Utilization of results may occur over a period of time; realistically the time period can vary according to individual characteristics of the program. Change is costly, and thus utilization, defined in terms of changes made in the program, may take both time and money. Utilization of research results may also require social change (Fairweather, 1977) and therefore consideration of the elements of a social change mechanism is necessary. Utilization can be a slow process, and in fact may require some level of expertise. Davis and Salasin (1975) suggest that perhaps evaluators should extend the range of their roles to encompass change consultation. In the ideal situation, where planners and evaluators are working closely together, this is likely to happen. There is also the possibility that utilization may commense prior to the final stages of evaluation research. According to Twain (1975) if utilization is carefully planned for and the users well informed, selected aspects of the research project can be applied, even when the research program continues.

An obvious but overlooked fact is that if decision makers are to

use the results of an evaluation, they must understand the results and the implications of the data (Mann and Likert, 1971). Communication between planners and evaluators will contribute to this understanding. It is therefore important to be aware of a "personal" factor in communication. Patton (1978) suggests that utilization of the data is dependent upon the interest, capabilities, and initiative of the individuals in charge of decision making and change. Translation, interpretation, meaning and relevance are established by persons who take the time to care; others will disregard the findings. Cultivating the necessary characteristics is therefore an important task for the evaluator.

The presentation of the research findings is also an important element in the utilization of research results. The report document is a piece of persuasive communication (Brown, Braskamp and Newman, 1978). Report style, communication theory and attribution theory should be considered in the development of a report which will transmit information. In the real situation, it is not uncommon for the evaluator to get caught up with his/her own style and forget the relevance of those for whom the data are intended. Patton (1978) suggests that the format, purpose and organization of the final report should be discussed and negotiated with decision makers and users. The final report may therefore become more understandable and intelligible with brief summaries, rather than one lengthy document. Patton also emphasizes that the final report is only one part of the process. There should be no surprises with the final report. The majority of information, data, results, etc. should have been discussed beforehand. Thus, with continual communication and interaction between planners and

evaluators, the final report is only a final summary of the evaluation research results.

As indicated in the process model flow chart (Figure 6), there is also consideration for the utilization of the results by the program evaluator, such as providing information to other evaluators regarding lessons learned, techniques employed, and a contribution to relevant theory in the topic area. The ethics of this situation can become rather involved regarding whether or not the evaluator has the right to disclose information about an evaluation for which he/she was paid. However, in the ideal situation, it is assumed that both evaluator and planner are interested in the advancement of science and program evaluation as a tool. Thus, disclosure of information could be worked out satisfactorily.

Cook (1978) suggests that there are three criteria by which evaluation research can be evaluated: utilization, knowledge building and institutionalization. The latter two relate to the possibilities of both planner and evaluator using the results of the evaluation. There have been some substantive gains in knowledge-building from evaluation research, such as in the area of mental health. The scientific aim is to accumulate a set of replicated findings that can be subsumed into some form of pattern, or concept, which may be applicable to other settings and other times. Therefore, evaluation research, like other research should be attempting to discover scientific facts, with as much consideration as possible for the scientific method of doing so. Institutionalization is the process of becoming integrated into the social structure of the science establishment. Cook (1978) concludes that evaluation research has become increasingly institutionalized over

the last decade. Thus, evaluation research can serve the scientist as well as the program planner.

B. THE REAL

1. Data Analysis

Analysis of Chicago Spacewatch data was a time consuming endeavor. I had expected it to take approximately two months; it took six. The element of time became a critical factor in the lack of utilization of results. Much of the time consuming aspects of the analysis were devoted to multivariate statistical analyses which were not of value to the program planners. The steps involved in data analysis consisted of the following:

- After data collection, data were prepared for computer entry by work study students who transferred data from questionnaires to obscan sheets.

- General descriptive statistics were run, using SPSS.

- A factor analysis was run on the items related to space awaremess, knowledge, attitude, etc. From this procedure, three factors were evident: space attitude, space knowledge and perceived benefits resulting from the space program.

- Correlations between these three variables were run, as well as correlations of these variables with demographic data.

- Differences between groups of respondents were run using the Multivar Statistical Program.

A good portion of the time spent on data analysis consisted of learning how to run the Multivar program. It is a difficult program, with insufficient documentation. It was, however, the best program available to adjust alpha levels according to the number of tests being

run; I was running several for correlations and tests of significant differences. Although these analyses took considerable time, I felt they were important, at that time. I had thought pre and post test differences would provide some evidence of the effects of Chicago Spacewatch; or confirm that there had been no effects in terms of changes in attitude, knowledge, etc.

This dedication to statistical analysis had some serious effects on the utilization of evaluation results. As a consequence of the sophisticated data analytic techniques my final reports were not completed until five months following program completion. By the time they were received, new projects were demanding the attention of the NSI staff. Only one person, to my knowledge, Tom Gorski, read the final reports and he left NSI shortly thereafter.

In addition, the data analysis which I felt was so extremely important, was of little use to the program planners. According to feedback from Gorski, descriptive information (such as the number of people attending certain functions) and the type of data which made Chicago Spacewatch "look good" were the focus of attention and selection and use.

Unfortunately, there was little communication between Gorski and myself during this period. As a consequence, there was a lack of agreed upon problem definitions and needs delineation between myself, the evaluator, and Gorski, the potential user of the research findings. This lack of communication presented serious problems. For example, shortly after the new year, I called Gorski to inform him that data analysis would, from that point on, take at least another month, and thus the final report was about six to eight weeks away (a conservative

estimate). He was not pleased, as he informed me that a Board of Directors meeting was scheduled within three weeks and he needed data and evaluation results. Had I been providing him with continual mini reports as data analysis proceeded, Gorski would have felt comfortable with meeting his directors and providing at least little bits of evaluation data. However, he had virtually nothing. Thus, to accomodate his request, I prepared a thirty page preliminary summary report of the data I had obtained to date, including whatever information I had processed to date. In retrospect, I believe that this report was used; whereas the subsequent final report submitted three months later was not. I think Gorski read the preliminary report and pulled out information of value to him for his presentation. Ironically, the multivariate data analysis was incomplete at that time.

Relevant comments from Gorski following the evaluation (provided on a post evaluation questionnaire I asked him to complete) add further support for my perceptions:

"My only criticism is one of time. A shorter period between the end of the data collection and final report would have perhaps sparked more interest from the powers that be. A further translation of what the results mean would be helpful as with any research - 99.9% of the world (including many high level personnel) do not understand one iota about research methods or terms. I am especially grateful we had such a dedicated and thorough person leading the evaluation. The Institute could have been hoodwinked so easily."

Major lessons learned in terms of the data analysis, as it related to the utilization of results include the following:

- Communication with the planner is extremely important. The evaluator needs to know what kind of data are desired, what type of report, and within what timeframe. This information should be communicated in the phase of developing the research design, not during

data analysis.

- Program planners and evaluators may be operating on different time frames. It is important for the evaluator to have the results and conclusions within the time frame specified by the planner if the results are to be utilized.

- Although this project was intended to be a learning experience, it is highly recommended that the evaluator not attempt to debug a new statistical program in the process of analyzing data. The Multivar program was so new and poorly documented, that even the author of the program was unable to solve some of the problems I encountered. This type of learning experience should not be at the expense of the program planners.

2. Results

A summary of the Chicago Spacewatch Evaluation results is provided in Appendix A, the Executive Summary of the final report. Regarding the overall program, theoretical relationships between attitudes toward the space program, knowledge of space related activity, etc. were tested. The results in this area support other similar research, that attitudes, knowledge, and awareness are positively correlated with one another. These results were of particular interest to me, yet seemingly of little interest to NSI. It was this type of data, however, that began to approach the foundational assumptions of NSI, and thus, were perhaps threatening and not easily dealt with. A major result of the evaluation was that NSI activities reached only people with positive attitudes toward space to begin with. There was little "awareness awakening" of the general public.

However, the educational programs and particularly the CMOS

program were successful in that the reviews by participants were good, attendance was high, and feedback was positive. However, it should be noted that CMOS was sponsored and coordinated by Chicago State University, not NSI.

Good information was provided by the Chicago Spacewatch evaluation which could have had a significant effect on future program planning of Spacewatch type programs. It was found that the educational programs provided the best focus and attention to the space program. These programs were well received by both students and teachers alike. Community programs were too diverse. The media programs and individual community programs were not effectively unified within Chicago Spacewatch. In addition, there was little evidence to support that there was any interest in the space program by the general public. Thus, Chicago Spacewatch was effective in reaching people who were already interested in the space program, and for whom space related activity is a salient issue.

To say that Chicago Spacewatch was a failure or success is difficult, however, for there were no criteria for success established prior to the evaluation. Judgement of attendance to programs was difficult without any indicators of what was expected. There were attendees to every program; however, in most cases it was my judgement that the attendance was poor in contrast to its potential. NSI's judgement was not always the same, however. Attempting to draw any conclusions regarding the effect of Chicago Spacewatch on those who did attend programs was also a difficult task. Since most respondents attending the programs showed an already existing pro-space attitude, any changes as a consequence of the program were not being measured by

the measuring instruments and data collection procedures. Furthermore, implementation was a problem, as discussed in the previous section. Chicago Spacewatch did not reach the general public and thus my pre and post surveys from the general public were not effective in assessing any impact. Although no differences could be found, this was attributed to the fact that Chicago Spacewatch did not reach the people.

With the large amount of data collected however, it was necessary to draw conclusions for NSI. On the one hand I knew it was my responsibility to provide an interpretation of the data, since it was unlikely NSI would be able to do so. However, on the other hand, I felt forced into making interpretive conclusions, especially those which would make NSI look good. I was reluctant to provide these bottom line type conclusions. As a consequence, NSI selected out data from the overall report that was best suited for their purposes.

The major lesson learned in this area is the critical importance of establishing criteria for success prior to program implementation. I was totally unaware of the expectations the NSI staff had in regards to Chicago Spacewatch. In lieu of their expectations, I developed those of my own, which were not within the same frame of reference as those of NSI. As a consequence, the wrong type of data were collected, conclusions were difficult to discern, and utilization of the results was negligible.

3. Utilization of Results

This is a difficult aspect of the evaluation to evaluate, in that I was in Chicago, and NSI is located in Virginia. Communication channels following program implementation were infrequent by telephone and mail until such time as the final report was delivered. Shortly

thereafter, all communication ceased. However, on the post program questionnaire which I sent Gorski, I asked him specifically to provide feedback regarding the utilization of the results. Although he answered many questions regarding this topic, most informative are the following comments:

"The evaluation was to be, from the start, a learning tool. When viewed in that perspective, precise methodology was not critical. The results were studied only by myself with a summary memo of lessons learned from the entire Spacewatch given to the executive director and executive committee. Excerpts of the findings were used when it was to the benefit of NSI--stressing the positive, such as the percentage of those who knew of Spacewatch or that something was going on at the Museum."

The major point regarding the utilization of results is that methodology and data analysis were critical to me as the evaluator; whereas Gorski used descriptive statistics and data which were supportive of NSI. Gorski left NSI shortly thereafter. To the best of my knowledge, the evaluation results have not been used by any other staff members.

Unfortunately, representatives from the Chicago Public School System, and Chicago State University, who had provided much information to me and who had also expressed interest in the results of the evaluation, were never provided any feedback. I did not feel I was in a position to release evaluation data; however, it is probable that they would have used the data in a constructive manner for future program planning. I asked Gorski if he would like a mini report prepared for these representatives. His response was that NSI would assume that responsibility. To my knowledge, this never came to pass.

Greater use of the data may have been facilitated if the results had been presented in person to the NSI staff in a briefing format to

ensure comprehension and understanding of the final results. There was little reception to such an idea, however. In retrospect, I also realize I should have sent several progress reports, as data analysis progressed, rather than just the preliminary report and final report. A little bit of information gradually presented, and continually presented in comprehensive language may have spurred greater utilization of the results. The valuable lesson learned is that the final report may not be read, and should be available as a reference document primarily.

It is also possible that NSI was unsure of how to apply the evaluation results in future program planning. Educating the NSI staff (particularly Gorski) about how to use the evaluation data may have had a positive influence on utilization. It is interesting to note, however, that with the exception of one small program in Las Vegas, NSI has not sponsored nor organized any other Chicago Spacewatch type programs since Chicago Spacewatch. There have been considerable political upheavals in the organization and a complete change-over in staff. Until such time as similar programs are initiated, evaluation results may not be utilized, as they are not needed.

4. Lessons Learned

Utilization of the results of an evaluation are one criterion of a good evaluation, one which provides the information necessary to assist program planners in decision making. Unfortunately, the evaluation of Chicago Spacewatch did not score high on utilization. The reasons for this are numerous, many of which began with problems in the earliest phases of the evaluation. The lessons learned through this phase are also reflective of lessons from previous phases:

- Good communication between planner and evaluator is essential.

- Criteria for success should be established prior to program implementation.

- The evaluator should be aware of differences in perspectives between planner and evaluator.

- The evaluator should be prepared for misuse of data; the planner will use the data in whatever manner is most useful for accomplishment of his/her goals and purpose.

I feel that lack of communication between myself and the NSI staff was the most serious problem and influential factor in the lack of evaluation utilization. In the ideal process model, good communication is established from the onset, with a proper understanding and appreciation for the other's role and perspective. In the case of Chicago Spacewatch, the earliest phases were bypassed due to a time crunch; the latter phases of the evaluation suffered as a consequence.

CHAPTER IX

CLOSING

In closing, the insight of Weiss and Rine (1972) of broad aimed programs seems relevant. According to these authors, evaluation of broad-aimed programs is plagued with technical difficulties, some of which are:

- It is difficult to select satisfactory criteria for success. There are many possible interpretations to the aims and goals of the program.
- The situation is essentially uncontrolled. Comparison cases are apt to be too few, non-random and potentially affected by the program itself.
- The research staff may know less than the action group about what is going on, rather than more.

All of these are true of Chicago Spacewatch and its evaluation. Chicago Spacewatch definitely qualifies as a broad-aimed program and the evaluation met with innumerable difficulties. Criteria for success were not established, causing major problems in developing the results and conclusions of the evaluation. The implementation of Chicago Spacewatch was a situation essentially uncontrolled. The program was too large and diverse for the small staff. The only comparison case used in the research design was a group of public school teachers who did not attend the programs. A selection bias was a potential problem in this regard. The research staff most assuredly knew less about the activities than the program staff. A good clear understanding of the program (and its

diversity) and the complicated pattern of sponsorship should have preceded the development of a research design.

However, Weiss and Rine (1972) suggest that negative circumstances and results should not be shrugged off, nor ignored; it is necessary for the research person to learn as much as possible from the experience, irrespective of its results. The research design should attempt to find the forces which shape the specification of the program and the nature of the opposition it encounters, reasons for failure and the program's unanticipated consequences. To some extent, this intention became the focus of the evaluation research after it became apparent that program implementation was a serious problem. Surveys were conducted in additional settings to assess advertising effectiveness and to determine how pro-space program attendants were. Attendance to meetings was monitored, and whenever possible, interviews were held with people to find out about their perceptions of program proceedings.

Most important, however, is that as the research person, I did in fact learn invaluable lessons from the experience. In addition to the more specific lessons learned, discussed in previous sections of this paper, general overall lessons include the following:

- The ideal process model should be adhered to in the best manner possible. The Chicago Spacewatch program and program evaluation could have been significantly improved if efforts to follow major steps in the model had been made. However, I had not developed the model prior to the evaluation. Its future use is intended.

- Communication and cooperation between planner and evaluator throughout each phase of the evaluation is of the utmost importance. Moreover the communication and cooperation must address the relevant

steps in the process model. In other words, both planners and evaluators should work together to define goals, clarify goals, identify criteria for success, negotiate program design and research designs, identify data needed and desired, how data will be used and in what format it should be presented. It may require a substantial amount of effort to ensure this type of relationship, yet any effort to eventuate the ideal can lead to a better program and a more useful evaluation.

● Be prepared for misunderstandings, expect failures and anticipate Murphy's law. The ideal process model is an ideal; the real may fall considerably short of it in several ways. The intention of the ideal is to provide a model to work towards; the attainment of the ideal is not realistic. Related to this point is the need for the program evaluator to be flexible, and adjust accordingly to the problems encountered.

The evaluation of Chicago Spacewatch was a valuable learning experience. I now have the experience and the lessons to accompany me in evaluation efforts of the future. With each successive evaluation, however, I intend to achieve a closer approximation of the ideal developed in this course of study.

REFERENCES

Agarwala-Rogers, R. (1977) Why is Evaluation Research Not Utilized? In M. Guttentag and S. Saar (Eds.) Evaluation Studies Review Annual Volume II. Beverly Hills: Sage Publications.

Alkin, M. C. (1972) Evaluation Theory Development. In C. H. Weiss (Ed.) Evaluating Action Programs. Boston: Allyn and Bacon Inc.

Alwin, D. F. and Sullivan, M. J. (1976) Issues of Design and Anlysis in Evaluation Research. In I. Bernstein (Ed.) Validity Issues in Evaluation Research. Beverly Hills: Sage Publications.

Andriew, M. (1977) Benefit cost evaluation. In L. Rutman Evaluation Research Methods: A Basic Guide. Beverly Hills: Sage Publications.

Berk, R. A. and Brewer, M. (1978) Feet of Clay in Nobnail Boots: An Assessment of Statistical Inference in Applied Research. In T. C. Cook Evaluation Studies Review Annual Volume 3. Beverly Hills: Sage Publications.

Bernstein, I. N.; Borhnstedt, G. W.; and Borgatta, E. F. (1976) External Validity and Evaluation Research. A codification of Problems. In I. Bernstein (Ed.) Validity Issues in Evaluation Research. Beverly Hills: Sage Publications.

Boruch, R. F. and Gomex, H. (1979) Sensitivity, Bias and Theory in Impact Evaluation. In L. Sechrest et al (Eds.) Evaluation Studies Review Annual Volume 4. Beverly Hills: Sage Publications.

Boruch, R. F. and Rindskoft, D. (1977) On randomized Experiments, Approximation to Experiments and Data Analysis. In L. Rutman (Ed.) Evaluation Research Methods: A Basic Guide. Beverly Hills: Sage Publications.

Brickell, H. M. (1978) The Influence of External Political Factors on the Role and Methodology of Evaluation. In T. C. Cook et al (Eds.) Evaluation Studies Review Annual Volume 3. Beverly Hills: Sage Publications.

Brown, R. D.; Broskamp, L. A.; and Newman, D. L. (1978) Evaluation Credibility as a Function of Report Style. Evaluation Quarterly, Volume 2, #2, May, 1978.

Cain, G. G. and Hillister, R. G. (1972) The Methodology of Evaluating Social Action Programs. In P. H. Rossi and W. Williams (Eds.) Evaluating Social Programs Theory, Practice, and Politics. New York: Seminar Press.

Campbell, D. T. (1975) Reforms as Experiments. In E. L. Struening and M. Guttentag (Eds.) Handbook of Evaluation Research Volume 1. Beverly Hills: Sage Publications.

Campbell, D. T. and Stanley, J. C. (1963) Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally Publishing Co.

Cherns, A. (1971) Social Research and its Diffusion. In F. Caro (Ed.) Readings in Evaluation Research. New York: Russell Sage Foundation.

Cohen, D. K. and Weiss, J. V. (1978) Social Science and Social Policy: Schools and Race. In T. C. Cook et al (Eds.) Evaluation Studies Review Annual Volume 3. Beverly Hills: Sage Publications.

Connor, R. F. (1978) Selecting a Control Group: An Analysis of the Randomization Process in Twelve Social Reform Programs. In T. C. Cook et al (Eds.) Evaluation Studies Review Annual Volume 3. Beverly Hills: Sage Publications.

Cook, T. D. (1978) Utilization, Knowledge-Building, and Institutionalization: Three Criteria by which Evaluation Research can be Evaluated. In T. C. Cook et al (Eds.) Evaluation Studies Review Annual Volume 3. Beverly Hills: Sage Publications.

Cook, T. D., Cook, F. L. and Mark, M. M. (1977) Randomized and Quasi-Experimental Designs in Evaluation Research. In L. Rutman (Ed.) Evaluation Research Methods: A Basic Guide. Beverly Hills: Sage Publications.

Cook, T. D. and Gruder, C. (1979) Metaevaluation Research. In Sechrest et al (Eds.) Evaluation Studies Review Annual Volume 4. Beverly Hills: Sage Publications.

Cook, T. D. and Pollard, W. E. (1977) Evaluation, Volume 4, pg. 161-4.

Davis, H. R. and Salasin, S. E. (1975) The Utilization of Evaluation. In E. L. Struening and M. Guttentag (Eds.) Handbook of Evaluation Research Volume 1. Beverly Hills: Sage Publications.

Deming, W. E. (1975) The Logic of Evaluation Research. In E. L. Struening and M. Guttentag (Eds.) Handbook of Evaluation Research Vodlume 1. Beverly Hills: Sage Publications

Eber, H. W. (1975) Mutivariate Methodologies for Evaluation Research. In E. L. Struening and M. Guttentag (Eds.) Handbook of Evaluation Research Volume 1. Beverly Hills: Sage Publications.

Edwards, W., Guttentag, M. and Snapper, K. (1975) A Decision - Theoretical Approach to Evaluation Research. In E. L. Struening and M. Guttentag (Eds.) Handbook of Evaluation Research Volume 1. Beverly Hills: Sage Publications.

Fairweather, G. W. (1977) A Process of Innovation and Dissemination Experimentation. In L. Rutman (Ed.) Evaluation Research Methods: A Basic Guide. Beverly Hills: Sage Publications.

Fishbein, M. and Ajzen, I. (1975) Belief, Attitude, Intention and

Behavior. Reading, Massachusetts: Addison-Wesley.

Freeman, H. E. (1977) The Present Status of Evaluation Research. In M. Guttentag and S. Saar (Eds.) Evaluation Studies Review Annual Volume 2. Beverly Hills: Sage Publications.

Graycar, A. (1979) Political Issues in Research and Evaluation. Evaluation Quarterly. Volume 3, Number 3, August 1979, pgs. 460-71.

Gurel, L. (1975) The Human Side of Evaluating Human Service Programs: Problems and Prospects. In M. Guttentag and E. Struening (Eds.) Handbook of Evaluation Research Volume 2. Beverly Hills: Sage Publications.

House, E. R. (1976) Justice in Evaluation. In G. V. Glass (Ed.) Evaluation Studies Review Annual Volume 1. Beverly Hills: Sage Publications.

LaPorte, T. and Metlay, D. (1975) They Watch and Wonder: Public Attitudes Toward Advanced Technology. Final Report to Ames Research Center National Aeronautics and Space Administration. NASA Grant NGR 05-003-0471. December 1975.

Larkey, P. D. (1979) Process Models of Government Resource Allocation and Program Evaluation. In L. Sechrest et al (Eds.) Evaluation Studies Review Annual Volume 4. Beverly Hills: Sage Publications.

Mann, F. and Likert, R. (1971) The Need for Research on the Communication of Research Results. In F. Caro (Ed.) Readings in Evaluation Research. New York: Russell Sage Foundation.

Nunnally, J. C. (1975) The Study of Change in Evaluation Research: Principles Concerning Measurement, Experimental Design, and Analysis. In E. L. Struening and M. Guttentag (Eds.) Handbook of Evaluation Research Volume 1. Beverly Hills: Sage Publications.

Nunnally, J. C. and Durham, R. L. (1975) Validity, Reliability and Special Problems of Measurement in Evaluation Research. In E. L. Struening and M. Guttentag (Eds.) Handbook of Evaluation Research Volume 1. Beverly Hills: Sage Publications.

Nunnally, J. C. and Wilson, W. H. (1975) Method and Theory for Developing Measures in Evaluation Research. In E. L. Struening and M. Guttentag (Eds.) Handbook of Evaluation Research Volume 1. Beverly Hills: Sage Publications.

Orata, P. T. (1940) Evaluating Evaluation. Journal of Education Research. Volume 33, pg. 641.

Patton, M. Q. (1978a) Utilization Focused Evaluation. Beverly Hills: Sage Publications.

Patton, M. Q.; Grimes, P. S.; Guthrie, K. M.; Brenna, N. J.; French, B. D.; Blyth, D. A. (1978) In Search of Impact: An Analysis of the

Utilization of Federal Health Evaluation Research. In T. C. Cook et al (Eds.) Evaluation Studies Review Annual Volume 3. Beverly Hills: Sage Publications.

Patton, M. Q. (1979) Evaluation of Program Implementation. In L. Sechrest et al (Eds.) Evaluation Studies Review Annual Volume 4. Beverly Hills: Sage Publications.

Rein, M. and White, S. (1978) Can Policy Research Help Policy? In T. C. Cook et al (Eds.) Evaluation Studies Review Annual Volume 3. Beverly Hills: Sage Publications.

Riecken, H. W. (1972) Program Evaluation. In Weiss, C. H. (Ed.) Evaluating Action Programs. Boston: Allyn and Bacon Inc.

Rossi, R. H., Freeman, H.W., and Wright, S. R. (1979) Evaluation: A Systematic Approach. Beverly Hills: Sage Publications.

Rossi, R. J. and McLaughlin, D. H. (1979) Establishing Objectives. Evaluation Quarterly. Volume 3, Number 3, pgs. 331-46.

Rossi, P. H. (1979) Issues in the Evaluation of Human Service Delivery. In L. Sechrest et al (Eds.) Evaluation Studies Review Annual Volume 4. Beverly Hills: Sage Publications.

Rutman, L. (1977) Formative Research and Program Evaluability. In L. Rutman (Ed.) Evaluation Research Methods: A Basic Guide. Beverly Hills: Sage Publications.

Scheirer, M. A. (1979) Program Participants' Positive Perceptions: Psychological Conflict of Interest in Social Program Evaluation. In L. Sechrest et al (Eds.) Evaluation Studies Review Annual Volume 4. Beverly Hills: Sage Publications.

Schulberg, H. C. and Baker, F. (1971) Program Evaluation Models and the Implementation of Research Findings. In F. Caro (Ed.) Readings in Evaluation Research. New York: Russell Sage Foundation.

Scriven, M. (1969) An Introduction to Meta-evaluation. Educational Product Report 2.

Scriven, M. (1976) Evaluation Bias and its Control. In G. V. Glass (Ed.) Evaluation Studies Review Annual Volume I. Beverly Hills: Sage Publications.

Sechrest, L., West, S.G., Phellips, R. R. and Yeaton, W. (1977) Some Neglected Problems in Evaluation Research: Strength and Integrity of Treatments. In L. Sechrest et al (Eds.) Evaluation Studies Review Annual Volume 4. Beverly Hills: Sage Publications.

Suchman, E. A. (1972) Action for What? In C. H. Weiss (Ed.) Evaluating Action Programs. Boston: Allyn and Bacon, Inc.

Tabor, J. C. (1978) The Role of the Accountant in Preventing and

Detecting Information Abuses in Social Program Evaluation. In T. C. Cook et al (Eds.) Evaluation Studies Review Annual Volume 3. Beverly Hills: Sage Publications.

Twain, D. (1975) Developing and Implementing a Research Strategy. In E. L. Struening and M. Guttentag (Eds.) Handbook of Evaluation Research Volume 1. Beverly Hills: Sage Publications.

Weiss, C. H. (1971) Utilization of Evaluation: Toward Comparative Study. In F. Caro (Ed.) Readings in Evaluation Research. New York: Russell Sage Foundation.

Weiss, C. H. (1972) Evaluation Research Methods of Assessing Program Effectiveness. Englewood Cliffs, N.J.: Prentice Hall, Inc.

Weiss, C. H. (1975) Evaluation Research in the Political Context. In E. L. Struening and M. Guttentag (Eds.) Handbook of Evaluation Research Volume 1. Beverly Hills: Sage Publications.

Weiss, R. S. and Rein, M. (1972) The Evaluation of Broad Aimed Programs: Difficulties in Experimental Design and an Alternative. In C. H. Weiss (Ed.) Evaluating Action Programs. Boston: Allyn and Bacon Inc.

Wholey, J. W., Scanlon, J. W., Duffy, H. G., Fukumoto, J. S., and Vojt, L. M. (1970) Federal Evaluation Policy. Washington D.C.: The Urban Institute.

Williams, W. (1971) Social Policy Research and Analysis. New York: American Elseviews Publishing Co.

Williams. W. (1978) Implimentation Analysis and Assessment. In T. C. Cook et al (Eds.) Evaluation Studies Review Annual Volume 3. Beverly Hills: Sage Publications.

APPENDIX A

EXECUTIVE SUMMARY
CHICAGO SPACEWATCH PROGRAM EVALUATION

PROGRAM DESCRIPTION

Chicago Spacewatch was a program of educational and community activites which took place in Chicago, Illinois, during the month of October, 1978.

The program began on October 1st, in observance of NASA's 20th anniversary. The program was designed to provide a linkage of information about space technology and space program activity to an awareness of how space technology and space related activity can be used as tools in solving problems here on earth.

The program was sponsored by the National Space Institute, NASA, Chicago Public Schools and Chicago Archdiocese Schools. In addition, many of the individual activites of Chicago Spacewatch were independently sponsored by local organizations.

The major role of the National Space Institute was to coordinate a series of programs within a month's duration which would provide repeated exposure to the public and educational institutions of the benefits of space applications to problems on earth.

PROGRAM ACTIVITIES

Media Programs

TV and radio programs were scheduled providing information about the space program, space technology and resulting benefits to people on earth.

Newspaper and magazine articles were also frequent during the month, similarly providing information about the benefits on earth of space activity.

Community Programs

Elderly Programs:

-NSI participation in Senior Citizen Day at St. James Cathedral (9/27/78)

-Space food adapted luncheon for the elderly at the Museum of Science and Industry (10/14/78)

Special Space Exhibits: (10/11/78 - 10/23/78)

-Museum of Science and Industry, including R2D2 and a Rockwell multi-media presentation

-Adler Planetarium

Celebrity Appearances:

-John Denver appearing at the Museum of Science and Industry (10/14/78)

-Gerald O'Neill, author of The High Frontier, appearing at Navy Pier (10/20/78)

Educational Programs

Teacher Orientation held at the Adler Planetarium (9/27/78)

School Speaker Program (9/25/78 - 10/25/78)

Chicago Meets Outer Space Program, held at Chicago State University (10/1/78 - 10/8/78)

Student contests (Jets, Getaway Special)

Teacher Workshop, held at Chicago State University (10/2 - 10/5/78)

High School Debates held at Navy Pier (10/20/78)

EVALUATION METHODOLOGY

The program evaluation of Chicago Spacewatch was arranged in the first week of September (1978). At this point, Chicago Spacewatch was an already designed and goal specified program. Therefore, the evaluation was designed according to the goals previously specified.

Due to the diversity of the individual programs and their respective goals within Chicago Spacewatch, the evaluation was divided into three phases, each phase determined by specific goals, population targets and activity.

Phase 1:

Phase 1 provides information regarding the overall program assessment. It is comprised of three separate sections.

The first section deals with the population of the public in general, investigating baseline measures of general space awareness and attitude from before Chicago Spacewatch to after, and public awareness of Chicago Spacewatch.

The second section deals with a population of local members of space related organizations, who are thought to be space enthusiasts. In order to verify the sensitivity of the instruments used in the evaluation to measure general space awareness and attitude, responses to measures of general space awareness and attitude by local members are compared to the responses by the general public.

In addition, information was obtained regarding member activities related to space and their future expectations.

The third section deals with the total respondent population in the evaluation of Chicago Spacewatch. Investigations in this phase are concerned with how knowledge of space related activity, benefits perceived as a result of the space program and an attitude toward the space program are related, as well as how demographic variables are related to these same measures.

Phase II:

Phase II provides information regarding the community programs assessment. It is comprised of several individual assessments of community programs.

Investigations include awareness of Chicago Spacewatch, characteristics of participants and the effect of participation in the programs on general space awareness and attitude.

Phase III:

Phase III provides information regarding the educational programs assessment. It is comprised of several individual assessments of educational programs.

Investigations include awareness of Chicago Spacewatch, characteristics of participants, teacher attitude toward teaching the study of space in the classroom, teacher attitude toward attending educational programs about space, measures of general space awareness and attitude, and the effectiveness of advertising and promotion.

The evaluation used a variety of methods for obtaining information, including the following:

-attendance records

-telephone interviews

-personal interviews

-questionnaires

For each phase and each respective individual program/section, one or several of these methods were used.

An overall description and general results of each phase of the evaluation are presented in the following sections of this report.

GENERAL CONCLUSIONS

Based on the overall evaluation of Chicago Spacewatch, several general conclusions have been drawn. They are as follows:

- Based on the reports of previous studies, it was expected that an attitude toward the space program would be highly correlated with knowledge of space related activity and benefits perceived as a result of the space program. The results of the evaluation confirm this expectation.

  It was found that persons with a positive attitude toward the space program are also more knowledgeable of space related activity and perceive more benefits resulting from the space program.

  Although this relationship was found to exist, it is not discernable from the results of the evaluation which is the cause or which is the effect.

- Chicago Spacewatch was effective in reaching people who were already interested in the space program and for whom space related activity is a salient issue.

Since most assessments of general space awareness and attitude were made prior to each individual program, the high level of interest in the space program cannot be considered an effect of Chicago Spacewatch, but rather should be considered the major characteristic of respondents to Chicago Spacewatch.

- The space program and resulting benefits are currently not a salient issue to the general public. Therefore, selective exposure to other newsworthy information precluded attention to Chicago Spacewatch and information dissemination.

  It is therefore recommended that community and media programs such as those of Chicago Spacewatch be implemented concurrently with some newsworthy issue of potential salience, such as the space shuttle flight.

- The media programs and individual community and educational programs were not effectively unified within Chicago Spacewatch.

  The majority of recipients in all programs were unaware of the overall program of Chicago Spacewatch, cognizant only of the particular programs attended. This is possibly a result of independent sponsorship for many programs.

- The educational programs were very successful in providing space education for many children during the course of Chicago Spacewatch. It is not evident, however, whether there will be a continuation of space education as a result of these programs.

- There is a need for educational programs which will assist teachers in integrating the study of space into the classroom.

  Teachers need information and materials and specific guidelines of how to incorporate space study into the classroom. However, acquisition of this information should be easy, convenient, condoned by the school board, and without excessive time and money demands of the teachers.

PHASE I: OVERALL PROGRAM ASSESSMENT

PHASE I

OVERALL PROGRAM ASSESSMENT

GENERAL DESCRIPTION

Phase I of the evaluation is an overall assessment of Chicago Spacewatch, and an overall assessment of the data collected during the course of the evaluation.

There are three separate sections in this phase, which are identified by respondent populations.

The first section is the general public population. Measures of general space awareness and attitude were obtained before, during and after Chicago Spacewatch.

The second section is the population of local members of space related organizations. Measures of general space awareness and attitude were obtained, as well as other information regarding space related activity and future expectations.

The third section is an analysis of responses by all persons who completed the basic questionnaire during the course of Chicago Spacewatch. Measures of general space awareness and attitude were obtained. Relationships between these measures are investigated, as well as the relationships between demographic variables and measures of general space awareness and attitude.

OVERALL PROGRAM GOAL

To stimulate public awareness, interest and understanding of space and the application of space technology as potential tools in solving problems on earth.

COURSES OF ACTION MANIFESTING THE GOAL

Community Programs: these programs are discussed in Phase II.

Education Programs: these programs are discussed in Phase III.

Media Programs: these programs included news releases to television, radio and newspapers providing information on upcoming events as well as information on the activities and benefits of the space program.

POPULATIONS FOR WHICH RESPONSES WERE ASSESSED

General Public

Local members of space related organizations

Total respondent population (all persons completing the basic questionnaire during the evaluation)

EVALUATION INSTRUMENTS AND CRITERIA

Telephone Interviews

A sample of 100 randomly selected Chicago residents were interviewed to determine the effectiveness of advertising and promotion.

Questionnaire

A basic questionnaire was designed to measure general space awareness and attitude to be used in all individual community and educational program evaluations, as well as the general public and local members of space related organizations.

The questionnaire used throughout the evaluation includes measures of the following:

- Awareness of Chicago Spacewatch
- General attitude toward the space program
- Knowledge of space related activity
- Benefits perceived as a result of the space program
- Heard of space program response
- Perceived income of the space program
- Reasons for space program continuation

OBJECTIVES OF EVALUATION PHASE I

The objectives of this phase of the evaluation were to provide information regarding the following:

-awareness of Chicago Spacewatch

-effectiveness of advertising and promotion

-baseline information regarding general space awareness and attitude of the general public

-change in general space awareness and attitude of the general public from before Chicago Spacewatch to after

-comparison of measures of general space awareness and attitude of the general public with those of local members of space related organizations

-participation in community programs and the effect on general space awareness and attitude

-correlation of measures of general space awareness and attitude with one another

-correlation of demographic characteristics (age, sex, income, education) with measures of general space awareness and attitude

GENERAL RESULTS AND CONCLUSIONS

General Public

The majority of the people interviewed had access to mediums of communication used in the dissemination of information in Chicago Spacewatch.

Respondents were most aware of R2D2 and the special space exhibits at the Museum.

In addition to the public media, friends and family were a common source of information about Chicago Spacewatch and community program activities.

Radio programs were most effective in providing specific information about the benefits of the space program.

Different people heard of different facets of the Chicago Spacewatch program, and it was not the same group who heard of each facet.

Of the questionnaire respondents who reported to have heard of Chicago Spacewatch after the program, TV was cited most frequently as the means of hearing of it.

Of 1,000 general public measured after Chicago Spacewatch, 14.1% reported attendance to one or several of the community programs.

The majority of respondents who attended community programs were not aware of Chicago Spacewatch.

The majority of the general public sampled was neither aware of Chicago Spacewatch nor received information being disseminated about the benefits from the space program.

There was no change in measures of general space awareness and attitude of the general public from before Chicago Spacewatch to after.

- Based on the low participation in community programs and the small response to hearing of Chicago Spacewatch, changes in space awareness and attitude were not expected.

  This suggests that the advertising and promotion of Chicago Spacewatch were ineffective in bringing people into the arena of activites, or there was no interest in the programs of Chicago Spacewatch and therefore people paid little or no attention.

- Information disseminated through public media channels was:

  - not effective in creating any change in space awareness or attitude

  - not effective in reaching the people

  - not of interest to the people and therefore they paid little or no attention.

    All of the above are considered likly possibilities.

Local members of space related organizations

Out of 126 members, 45.2% reported hearing of Chicago Spacewatch.

Out of 126 members, 42.7% reported attendance to special space programs during October.

The member population obtained the highest, most positive scores on all measures of general space awareness and attitude of all populations measured during the course of Chicago Spacewatch.

Members consider themselves more aware, knowledgeable, and interested in space related activity than the average citizen.

In general, members indicate a moderate amount of time devoted to space related activity.

Although members indicate it is important for the general public to be aware of space related activity, they indicate few intentions to participate in educating the public.

A possible reason for the lack of participation in educating the public is that members indicate they do not feel qualified to give lectures or lead discussions.

Regarding expectations of future space activity, members indicate the following:

- to a small extent, members feel that their future employment will be related to the space industry.
- to some extent members believe that space colonization will occur in their lifetime.
- members do not believe strongly that they will be space travelers.

According to members, research and development are the most important reasons for moving out into space.

Total Respondent Population

The overall theme of Chicago Spacewatch was not known by the majority of all respondent populations.

People for whom the space program is a salient issue heard of Chicago Spacewatch the most.

Overall, the newspaper was cited most frequently as the means of hearing of Chicago Spacewatch.

Affiliation with local space related organizations and the Chicago Public School System were also cited as a means of hearing of Chicago Spacewatch.

Local members of space related organizations had the highest percentage of respondents who attended programs during Chicago Spacewatch.

Participating teachers in the educational programs attended community programs more so than teachers who did not attend any educational programs.

In general, respondents have a more positive than negative attitude toward the space program.

People for whom space issues are salient and of interest, have the most positive attitude toward the space program and are most knowledgeable of space related activity.

Program participants, as compared to non-participants obtained higher, more positive scores on all measures of general space awareness and attitude.

Discovery of new energy and material resources is considered to be the most important reason for continuing the space program. Using space as a tool to solve problems on earth is considered the second most important reason.

Middle age respondents (age 26 - 50) evidence the most awareness of the space program and the most positive attitude toward the space program.

Males are more aware of the space program and have a more positive attitude than females.

The higher the level of income, the more awareness of the space program and the more positive attitude toward the space program is evidenced.

The higher the level of education, the more awareness and the most positive attitude toward the space program is evidenced.

A linear relationship was found between all measures of general space awareness and attitude.

> Therefore, persons with a positive attitude toward the space program are more knowledgeable of space related activity, perceive more benefits resulting from the space program, are more accurate in their perception of the budget of the space program, and have heard of the space program more.
>
> Although these relationships were found to exist, it is not discernable which are the causes or which are the effects.

PHASE II: COMMUNITY PROGRAMS ASSESSMENT

PHASE II

COMMUNITY PROGRAMS ASSESSMENT

GENERAL DESCRIPTION

Phase II of the evaluation consists of an evaluation of several individual community programs of Chicago Spacewatch.

COMMUNITY PROGRAM GOALS

1) To stimulate public awareness of space and space related technology to solutions of earth's problems of energy, environment, employment, food, etc.

2) To show the relevance of space to various aspects of life including the humanities, alternative lifestyles, etc.

3) To stimulate community cooperation and communication in space related programs.

COURSES OF ACTION MANIFESTING THE GOALS

-Disseminate information about space related activity to the general public via TV and radio programs, newspaper and magazine articles, exhibits, and lectures.

-Provide programs demonstrating the relevance to various aspects of life.

-Advertise and promote attendance to exhibits, lectures and programs.

INDIVIDUAL PROGRAMS EVALUATED

Elderly Programs

1) NSI participation in Senior Citizen Day at St. James Cathedral
2) Space adapted food luncheon for the elderly at the Museum of Science and Industry

Museum of Science and Industry

Including: special space exhibits, celebrity appearance of John Denver, R2D2 and the Rockwell multi-media presentation.

O'Neill Presentation

EVALUATION INSTRUMENTS AND CRITERIA

Attendance

Attendance records are presented for each program. In some cases approximations are used where accurate data are unavailable.

Personal Interviews

Personal interviews were used in the evaluation of the elderly programs.

Questionnaire

A questionnaire was used in the evaluation of the O'Neill presentation and the Museum of Science and Industry programs. It includes measures of the following:

- Awareness of Chicago Spacewatch
- General attitude toward the space program
- Knowledge of space related activity
- Benefits perceived as a result of the space program
- Heard of space program response
- Perceived income of the space program

OBJECTIVES OF THE EVALUATION PHASE II

The objectives of this phase of the evaluation were to provide information regarding the following:

-program attendance and participation

-awareness of Chicago Spacewatch

-advertising effectiveness

-effect of programs on general space awareness and attitude

GENERAL RESULTS AND CONCLUSIONS

Elderly Programs

NSI Participation in Senior Citizen Day

Several hundred elderly and handicapped individuals attended this function.

The main reason given for attendance was a free box lunch and entertainment. All interviewees were surprised by questions related to space and inquired of their relevance.

None of the elderly interviewed were aware of the space exhibits set up in the front lobby, although all had entered through the front door.

The majority of those sampling astronaut food were unaware of what they were eating, many believing it was the free food they had been promised.

The speech given about the benefits resulting from the space program seemed to have very little effect on the elderly.

Space Adapted Food Luncheon

Approximately 200 elderly attended this luncheon.

The program was effective in stimulating awareness and appreciation of the space program; however, this occurred as a result of the advertising and promotion of the program through local nutrition centers, rather than the program itself.

The elderly were cognizant of the focus on the space program and aware that they were eating space adapted food.

The elderly were in favor of the space program before the program began, in addition to being aware of many benefits of the space program.

All elderly interviewed after the program indicated they had enjoyed the program.

In general:

The elderly feel that there have been benefits to the general public from the space program; however, in most cases they are unable to specify what these benefits are.

The elderly are unaware of personal benefits resulting from the space program.

The majority of the elderly think the space program should be continued, indicating national security as the major reason.

Museum of Science and Industry

During Chicago Spacewatch, Museum attendance was 35% higher than the average of the two preceeding years of comparable time periods.

On the weekend of October 14th and 15th, Museum attendance was 91% higher than the average of the two preceeding years for comparable weekends.

Advertisement of Chicago Spacewatch activities is considered to be a major factor in the increased attendance to the Museum during the period of Chicago Spacewatch.

The majority of museum participants had not heard of Chicago Spacewatch.

People attending for reasons of Chicago Spacewatch activities and those observed viewing special space exhibits were more aware of Chicago Spacewatch than the rest of the Museum respondent population.

The most frequently cited medium for hearing of Chicago Spacewatch was the newspaper.

In most cases, Chicago Spacewatch activites were not connected with the overall program of Chicago Spacewatch.

Respondents attending the Museum for reasons of Chicago Spacewatch activites were already in favor of the space program and did not change in space awareness and attitude as a result of their participation.

O'Neill Presentation

Approximately 500 people attended the presentation.

The majority of the respondent population reported having heard of Chicago Spacewatch, citing most frequently their local affiliation with a local space related organization as the means of hearing of it.

This respondent population reported hearing of Chicago Spacewatch more so than any other population measured during Chicago Spacewatch.

This respondent population evidenced an existing pro-attitude toward the space program, which is considered to be a primary reason for attendance to the presentation.

This respondent population had the highest, most positive scores on all measures of general space awareness and attitude of all populations measured during Chicago Spacewatch, with the exception of members of local space organizations, with whom responses were very comparable.

PHASE III: EDUCATIONAL PROGRAMS ASSESSMENT

PHASE III

EDUCATIONAL PROGRAMS ASSESSMENT

GENERAL DESCRIPTION

Phase III of the evaluation consists of an evaluation of the educational programs of Chicago Spacewatch. Some of the programs were directed towards teachers only and some to both teachers and students.

Many of the programs were independently organized through local institutions and organizations in cooperation with the National Space Institute.

EDUCATIONAL PROGRAM GOALS

1) To encourage student involvement in activites and projects to further understanding of space benefits and concepts, and to stimulate awareness of the application of space related technological developments in solving the earth's problems of energy, environment, food, population, employment, etc.

2) To integrate space concepts into all subject areas including the humanities (art, literature, etc.) by creating classroom and school projects, programs and activites that compliment existing curriculum.

3) To discuss the potentials of space, including career opportunities, benefits, alternative lifestyles, etc.

4) To give special recognition to outstanding individuals and projects.

COURSES OF ACTION MANIFESTING THE GOALS

-Provide students with programs of discussion, lectures, movies, etc. of space related activity.

-Provide teachers with information and ideas for incorporating the study of space related subjects into the classroom.

-Promote student contests related to space subjects (i.e. essay contest, Jets contest, Getaway Special).

INDIVIDUAL PROGRAMS EVALUATED

Teacher Orientation

School Speaker Program

Chicago Meets Outer Space Program

Teacher Workshop

Jets Student Contest

EVALUATION INSTRUMENTS AND CRITERIA

Attendance

Attendance records are presented for each program. In some cases approximations are used where accurate data are unavailable.

Interviews

Interviews were conducted with program coordinators and speakers.

Questionnaire

A questionnaire was designed for the teachers that participated in each of the programs. A basic questionnaire was designed for all teachers; however, the teachers in each program received a questionnaire modified with respect to the particular program.

The basic questionnaire was designed to obtain the following:

- A rating of the program on several characteristics, such as informative and relevant for classes.
- A measure of attitude about attending such programs and considerations for attending similar future programs.
- A measure of attitude regarding the study of space in the classroom.
- A measure of general space awareness and attitude

OBJECTIVES OF THE EVALUATION PHASE III

The objectives of this phase of the evaluation were to provide information regarding the following:

- program attendance and utilization
- participating teacher ratings of the educational programs
- teacher attitude toward teaching the study of space in the classroom
- teacher attitude toward attending educational programs
- general space awareness and attitude of teachers
- effect of educational programs on general space awareness and attitude

GENERAL RESULTS AND CONCLUSIONS

Attendance

Teacher Orientation: approximately 200 teachers attended the orientation; 76 schools are known to have been represented.

School Speaker Program: 35 schools requested a speaker; the target goal was 100.

Chicago Meets Outer Space: 61 schools were represented; over 4,000 children attended.

Teacher Workshop: 8 teachers attended.

Jets Contest: 20 applications were received; however, only 3 teams attended the contest.

Characteristics of partipating teachers

Educational programs were attended primarily by experienced upper-grade level science teachers.

Social studies teachers were very poorly represented in all programs. Ineffective advertising and lack of interest are two possible reasons for this under representation.

Awareness of Chicago Spacewatch

The overall theme of Chicago Spacewatch was not known by the majority of teachers.

For those who heard of Chicago Spacewatch, the most frequently cited means of hearing of it was the newspaper and the individual educational program announcements/bulletins.

The majority of teachers did not attend the community programs.

Program advertising

Personal invitations extended by mail or telephone resulted in the best response to the programs. Letters to teachers and school bulletins were also effective.

Not all science teachers were reached in the advertising campaign, even though this group was the target of most advertisements.

Time and location of the programs were important considerations in attendance by teachers. Time and location were good for those who did attend, and not good for those who did not.

As a consequence of individual advertising for each program, some teachers received several announcements, one for each program, whereas others received none. Workshop teachers reported this overlap in advertising to be confusing.

Program ratings

All programs were rated positively by participating teachers.

Chicago Meets Outer Space was rated most relevant to education.

All programs were rated interesting and informative.

Ratings on useful for classes and relevant to my concerns (teachers' concerns) were not high.

The lowest ratings were obtained on "organized" on all programs.

Teacher attitude toward attending space educational programs

Teachers in all programs report that not much time has been devoted to the study of space in the classroom as a result of the programs.

Having attended these programs, suggests there is a greater liklihood of attending future programs.

School board approval and positive consequences are expected from attending future programs.

It is more likely that science teachers will attend than social studies teachers.

Teacher attitude toward the study of space in the classroom

Not much time has been devoted to the study of space in the classroom during the 77/78 school year.

More time would be considered with information and materials.

Students have shown some interest in the study of space.

Teachers indicate students are unaware of a career potential in space and think it is important for students to have this awareness.

Teachers report they would enjoy the study of space in the classroom and think their students would also.

Teachers think it is important for other teachers to include the study of space in the classroom.

Teachers think it is important for education al all levels from elementary to college, to address the study of space.

Teachers report some difficulty in including the study of space into current teaching programs.

As a result of attending the teacher orientation, teachers evidenced a more positive attitude toward teaching space in the classroom.

Teachers attending the Chicago Meets Outer Space Program have the most positive attitude overall toward teaching space study in the classroom.

Participating teachers have a more positive attitude toward the study of space in the classroom, than a comparison group of teachers who did not participate.

General space awareness and attitude

Compared to the general public, all teachers evidence a more positive attitude toward the space program, more knowledge of space related activity, more benefits perceived as a result of the space program, more accurate perception of the income of the space program and have heard of the space program more.

As a result of the Teacher Orientation, participating teachers perceived more benefits resulting from the space program.

APPROVAL SHEET

The thesis submitted by Sherry R. McNeal has been read and approved by the following committee:

Dr. John Edwards, Director
Associate Professor, Psychology, Loyola

Dr. Jill N. Nagy
Associate Professor, Psychology, Loyola

The final copies have been examined by the director of the thesis and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the thesis is now given final approval by the Committee with reference to content and form.

The thesis is therefore accepted in partial fulfillment of the requirements for the degree of Master of Arts.

September 22, 1981
Date

John O. Edwards
Signature