HIGHWAY SAFETY PROJECT
EVALUATION METHODOLOGIES

FINAL REPORT

15 May, 1968 - 31 December, 1968

William Carlson
William Hall

31 December, 1968

Prepared by
HIGHWAY SAFETY RESEARCH INSTITUTE
Institute of Science and Technology
The University of Michigan
Ann Arbor, 48108

CONTENTS

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.
HSRI-SYA-2

2. Government Accession No.

3. Recipient's Catalog No.

4. Title and Subtitle
Highway Safety Project Evaluation Methodologies

5. Report Date
31 December 1968

6. Performing Organization Code

7. Author(s)
W. Carlson, W. Hall

8. Performing Organization Report No.

9. Performing Organization Name and Address
Highway Safety Research Institute
The University of Michigan
Huron Parkway & Baxter Road
Ann Arbor, Michigan 48105

10. Work Unit No.

11. Contract or Grant No.

12. Sponsoring Agency Name and Address
Michigan State Office of Highway Safety Planning
Lansing, Michigan

13. Type of Report and Period Covered
Final report
15 May 1968 - 31 Dec.1968

14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

The objective of this project was to develop and present techniques which would be useful in the evaluation of highway safety countermeasures programs conducted by local and state agencies. Systems analysis techniques were used to organize the evaluation process and to develop its components in a logical manner. The report includes a chronological overview of the program formulation, implementation and evaluation process, program modeling techniques, material on experimental design, cost-effectiveness analysis, and the application of these evaluative methods to four OHSP programs.

17. Key Words
Highway Safety Countermeasures
Evaluation

18. Distribution Statement
Unlimited

19. Security Classif. (of this report)
Unclassified

20. Security Classif. (of this page)
Unclassified

21. No. of Pages
171

22. Price

Form DOT F 1700.7 (8-69)

ACKNOWLEDGMENTS

The authors would like to thank Mr. Noel Bufe and the staff of the Office of Highway Safety Planning for the assistance which they provided during this study. Various Livonia City officials, Mr. Dean Wilson and Dr. William Robertson of the Pontiac School System, Captain John Plants and Mr. John Longstreth of the Michigan State Police, and Mr. Douglas Savala and Mr. Charles Durocher of the Michigan Department of State provided information and cooperation in the preparation of the applications section of this report. Finally, the authors would like to acknowledge the valuable assistance provided by Mr. James O'Day and Mr. Lyle Filkins of the Highway Safety Research Institute.

FIGURES

INTRODUCTION

This final report covers the work conducted from 15 May, 1968-31 December, 1968 on Highway Safety Project Evaluation Methodologies, a contract between the Office of Highway Safety Planning and the Highway Safety Research Institute of The University of Michigan.

The complete evaluation of highway safety projects is a difficult and complex task, requiring technical expertise, knowledge of the "state of the art", judgment, and careful analysis. Consequently, there is no one procedure or set of procedures which can be routinely used in evaluation. Instead the highway safety community must have flexible tools and methods which can be applied to the evaluation of projects.

The objective of this contract is to develop and present techniques which are useful in the evaluation of "action" projects conducted by local and state agencies. Throughout this study, we have used the techniques of systems analysis to organize the evaluation process and to develop the components of this process in a logical manner. Without such a systems approach, the evaluator is left with no unified way to consider project proposals.

In Chapter 1 we present an overview of the evaluation process by considering the chronology of events which should take place in project formulation and implementation. This chronology allows one to identify the necessary tasks at each step of the evaluation process and guides the execution of these steps.

Chapter 2 is devoted to a discussion of the early steps in evaluation--systematic classification and analysis. This procedure is necessary to isolate potential project effects and to discover those uncontrolled variables which might mask or exaggerate such effects.

The more technical aspects of this analysis are discussed in Chapter 3, where project modeling is discussed in some detail. By developing such a project model, the staff of OHSP will be able to select intermediate measures of project effectiveness. The resulting project evaluation criteria will then be meaningful in terms of the "causal chain" relating the project to a change in the collision process.

In certain situations an experiment may be conducted as a part of the project execution. Chapters 4 and 5 present material on experimental design which the staff of OHSP should understand--both in deciding whether to experiment and in developing meaningful experiments.

Cost-effectiveness analysis is considered in Chapter 6. In this chapter we discuss the essential components of cost-effectiveness analysis and the problems in applying standard cost-benefit ratios to highway safety projects.

The applications of these basic concepts to actual projects are presented in Chapters 7-10. These applications provide:

(1) Further explanation of the basic concepts developed in Chapters 1-6.

(2) An evaluation of applicability of these concepts to "real" projects.

(3) An independent analysis of four projects which are currently being considered by OHSP.

The projects which are chosen by OHSP display problems which fall into different portions of the chronology developed in Chapter 1. Consequently, their analysis is conducted in terms of this chronology. References are made to the basic material of the earlier chapters in order to integrate and expand upon these sections.

RECOMMENDATIONS

1. Problem Definition/Project Development

    i. OHSP should be require agencies to define their problems with more precision. When an agency is not able to do this, OHSP should encourage it to obtain the services of a competent consulting firm for a problem definition study.

    ii. OHSP should require agencies to develop a project by defining project objectives, components, and potential alternative projects. An information requirement table (Figure 2) will be useful in this effort.

2. A Priori Evaluation

    i. OHSP should determine from the proposal submitted, and if necessary from discussions with the agency, whether the agency has defined and developed its project effectively. If this has not been done, OHSP should return the proposal to the agency for further information. A possible alternative is for OHSP to assist the agency in defining and developing the project.

    ii. OHSP should consider whether the stated objective of the project is useful in relation to the overall safety objectives of the agency and/or OHSP.

iii. OHSP should classify projects and develop project models to determine potential project effects and potential biasing variables. This analysis should be used to select project evaluation criteria.

iv. OHSP should estimate the likelihood that the project, if properly executed, will achieve the chosen objectives. This estimate should be combined with an evaluation of the agency's ability to execute the project, thereby developing a subjective estimate of the overall project worth. This should be combined with cost information to develop a subjective project rating.

3. Project Conduct

i. OHSP should encourage extensive project planning by the agency after the contract award. This may be done by requiring a planning report or briefing early in the contract period.

ii. OHSP should consider an experimental evaluation if:

a) Such experimental information is unavailable and is beneficial to the agency and/or OHSP is evaluating this project or in planning future programs.

b) An experiment can be designed to obtain the desired information in a cost-effective way.

iii. If an experimental evaluation is part of the project, OHSP should insist that the experiment be well-designed and executed. This may be accomplished by:

a) The use of an agency staff member who is knowledgeable in the application of experimental design techniques.

b) The active participation of an OHSP staff member who is familiar with the project and the material of Chapters 4--Overview of Experimental Design--and 5--Experimental Considerations Particular to Highway Safety Project Evaluation.

c) Encouraging the agencies to obtain the services of an independent research unit in the experimental design and analysis.

4. Ex Post Facto Evaluation

i. OHSP should conduct an ex post facto evaluation by comparing the agency's planned objectives and costs with those attained. For future planning, an analysis should be made to determine the causes of discrepancies between the plan and the actual execution.

ii. Additional benefits which were unplanned but achieved should be noted for future planning.

5. General

i. OHSP should determine gaps in the "causal chain" relating a project to overall system effectiveness and should encourage research directed to filling these gaps in order to promote better future evaluations.

ii. OHSP should have the relevant material from this report digested into a reference manual for distribution to local and state agencies.

1. OVERVIEW OF THE EVALUATION PROCESS

In this chapter we wish to introduce and discuss the general concepts and problems of highway safety project evaluation. We shall do this by considering the entire chronology of project formulation and implementation. As will be indicated, project evaluation is an integral part of this chronology.

Several benefits come from looking at projects in this way. Specifically, such an overview provides guidance in answering the following general questions which have been posed by the staff of OHSP at various times during this study:

(1) Why should projects be evaluated?
(2) When should projects be evaluated?
(3) Who should evaluate projects?
(4) How does a systems "model" help in project evaluation?
(5) When should an experimental evaluation be conducted?
(6) How should the experiment be conducted?
    (a) Is there a standard checklist?
    (b) Is there a standard "model"?

Basically, a project can be divided into five sequential stages: problem definition, project development, a priori project evaluation, project conduct, and ex post facto project evaluation. The problem definition and project development stages will generally be undertaken by the agency which is submitting a proposal. The a priori evaluation will be done by OHSP. The project will be conducted by the agency with OHSP support. Finally, the ex post facto evaluation will be undertaken by OHSP.

These stages are presented as Figure 1. Within each stage there are a set of logically interrelated actions and questions which must be examined. We now consider these in greater detail.

PROBLEM DEFINITION

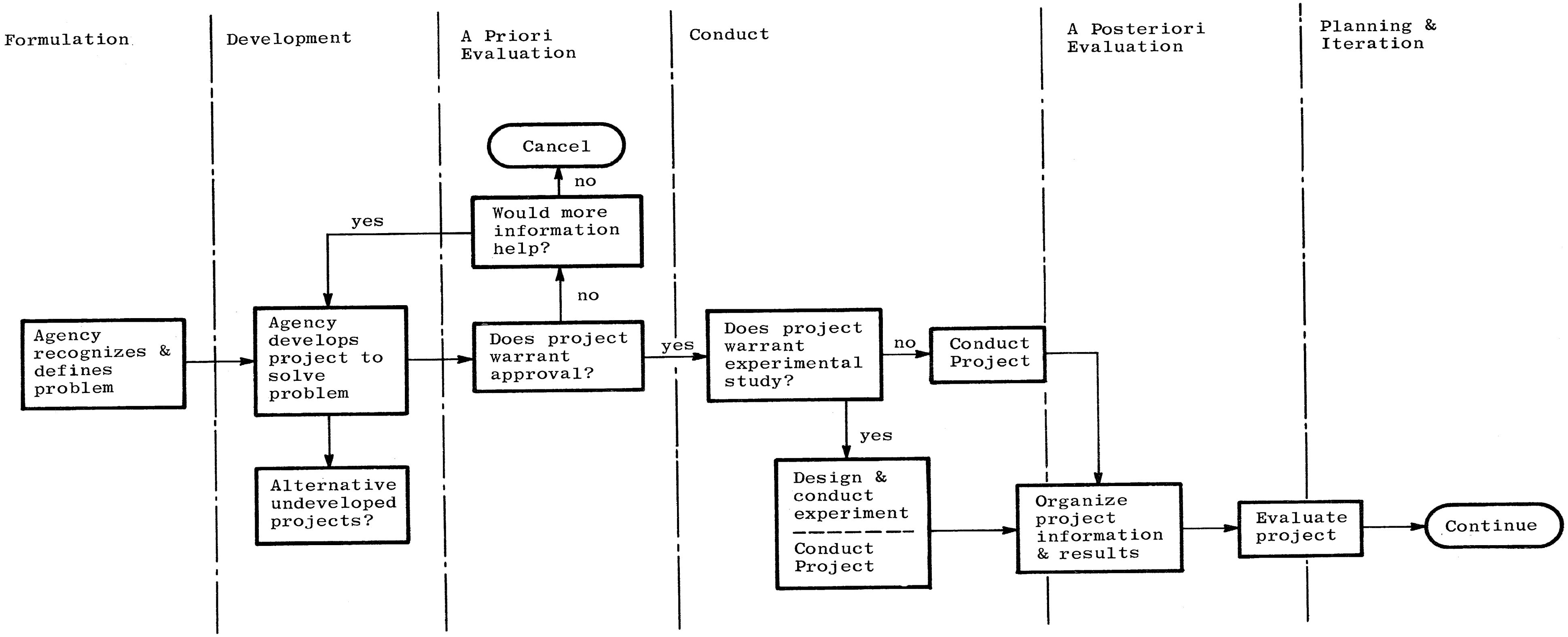
The first step in initiating a project consists of problem recognition and definition. This step is generally handled by the agency which will later develop a specific project designed to solve this problem. The proposal to do this will then be sumitted to OHSP.

The importance of clear and accurate problem definition on the part of the agency cannot be overstressed, for the failure to adequately perform this task may result in the "right solution to the wrong problem".

Problem definition may proceed at either of two levels. The agency may recognize that a problem exists; however, the agency also recognizes that it does not have adequate knowledge or ability to formulate (and eventually develop a project to solve) this problem. In this case the agency needs to undertake a problem definition study to determine the nature and extent of the problem. Such a project is designed to provide

Figure 1

CHRONOLOGY OF PROJECT FORMULATION AND EVALUATION



information useful in developing future action projects. In such a situation it is essential for the agency to assess as completely as possible the nature and extent of its lack of knowledge about the problem area. In this way the objectives of the definitional study may be specified prior to the start of the study.

An example of such a problem is that faced by the City of Livonia (Chapter 7). In this case the city wishes to uncover (forecast) the environmental problems (streets, signs, flows, etc) which will occur over the next twenty years. Here the problem definition should consist of a detailed consideration of the nature and extent of the uncertainty about these future events.

On the other hand, the local agency may explicity recognize the general nature of the problem. In this case problem definition will consist of careful and accurate determination of the nature and extent of the problem. In this way the agency can organize sub-projects to handle all components of the problem.

For example, consider the problem faced by Pontiac schools (Chapter 8). Here the problem has been recognized--the young driver is a problem; a new high school is under construction; hence how can we (Pontiac) develop a new program of driver education to improve this class of operators? Further exploration of the nature of the problem has revealed, for instance, that 1/2 of all juvenile court cases involve stolen cars. Hence the problem involves more than improved driver training alone. Similarly the problem involves a larger set of youths than those taking driver education to get a license. Further exploration of the extent of the problem has revealed that a comprehensive program for all school-aged children is desirable.

PROJECT DEVELOPMENT

After a problem has been defined, the local agency will develop a project to solve (or alleviate) this problem. In order to do this it is essential for the local agency to develop an explicit set of project objectives. This set of objectives should consist of overall general objectives as well as specific intermediate objectives which relate to these overall objectives. Specific projects can then be thought of as means of reaching these intermediate objectives.

In order for the local agency to correctly select a specific project, it should first consider whether there are reasonable alternative projects which might achieve the objectives delineated above. Alternative projects can then be compared by asking the following questions:

(1) What subset of the ojectives does each project affect? What is the likelihood that the project will achieve those objectives that it affects?

(2) What portion of the problem population is affected

by each project? In other words, what is the extent of the project?

(3) What are the costs, labor, and time involved in carrying out alternative projects? Are these feasible with current community resources and community support?*

This procedure results in the selection of a project which best satisfies the agency objectives within the given constraints. The accepted project is then developed by detailing potential benefits, techniques, costs, and personnel. The rejected projects are obviously left undeveloped.

Although the development of a project is primarily an agency function, there are obvious important implications from the point of view of OHSP. By carefully and explicitly considering objectives and alternative means of achieving these objectives, the agency is much more likely to develop a successful project. Therefore, it seems reasonable for OHSP to examine methods for encouraging such planning. One way to do this is to issue a directive to relevant agencies expanding upon these ideas. A second is to require in the proposal a specific and detailed listing of those objectives, sub-objectives, and any alternatives which have been analyzed by the local agency. This information would obviously differ for different projects, even for projects within the same

---

*The importance of such community support is stressed in References 1 and 2. In addition, these references provide a good overview of the highway safety program planning process.

functional area. However, a framework for such a presentation can be developed. One such outline is presented as Figure 2.

This outline can serve as one portion of the information required to be submitted with each project proposal. This outline will help guide agency development and in addition, it provides OHSP with a short summary of the agency's perception and analysis of the problem. As such the outline is one source of information for the subsequent evaluation.

A PRIORI EVALUATION

At this stage the basic decision is whether to conduct the project or not. We view this primarily as a function of OHSP.

Four basic sets of questions must be answered in making this decision:

(1) Has the agency carefully defined its problem and developed a project with specific objectives? Is the stated objective of value in improving highway safety?

(2) Is the project conceptually effective? (i.e., is the project capable of reaching its objectives if implemented correctly?)

(3) Can the agency effectively implement the project?

(4) Can the project be successfully completed in a cost-effective manner?

Although these questions are simple to ask we all agree that they are difficult to answer. Therefore we shall spend considerable time in developing systematic ways to analyze their ramifications.

FIGURE 2

EXAMPLE OF A PLANNING SHEET FOR FORMULATION & DEVELOPMENT

I. Problem Definition
  A. General statement as to nature of problem

  B. General statement as to extent (in terms of severity) populations affected.

II. Project Development
  A. General statement of agency objectives in attacking problem

  B. Statement of specific intermediate objectives to be achieved in reaching IIA

  C. Alternative projects considered for achieving IIA and IIB

    (1) Project 1

      a) Specific intermediate objectives affected
      b) Likelihood/magnitude of change in these associated with project
      c) % of each problem population affected by project
      d) Cost, labor, time analysis
      e) Potential problems in implementation
      f) Spin-off effects on community sub-groups (benefits/cost)

    (2) Project 2

      (a-f) above

  D. Analysis-reasoning for selection of specific project

  E. Detailed development of specific project chosen in IID

Throughout this study, HSRI has argued that systems analysis techniques can be applied to develop conceptual models of the projects OHSP is trying to evaluate. By "conceptual models", we have meant a hierarchical set of simplified causal relations specifying intermediate and final effects on the highway safety system.

We now propose to develop another model--this time of the a priori evaluation stage of Figure 1. Our intent in developing this model of an ideal evaluation is to organize this process as much as possible and to provide a systematic way of utilizing all relevant information in making a decision. The model is presented in Figure 3. Observe that the structure is no more than an ordered set of "boxes" indicating explicit information usage, comparisions, and actions.

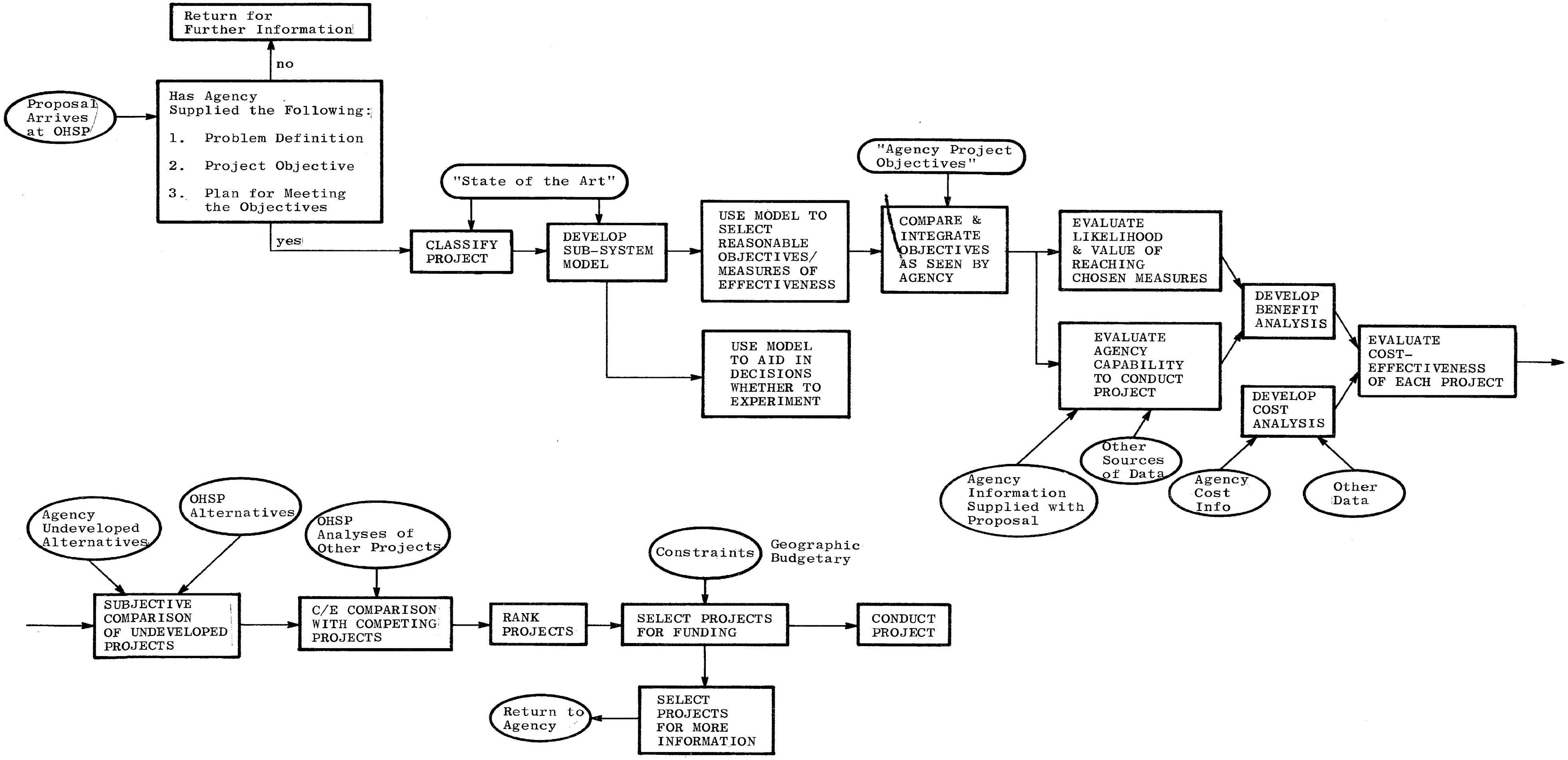
An essential first step in project evaluation is the classification of each project within a functional area into a set of meaningful categories. The development of such a classification system is considered in detail in Chapter 2. By utilizing this system the staff of OHSP can quickly ascertain the relevant components and phases of the overall highway safety problem which are affected by the project. In addition the type of action proposed by the project will also be determined by such an analysis.

At this stage a model of the potential project effects can be developed. Such a model will consist of logically

Figure 3

CONCEPTUALIZATION OF THE A PRIORI EVALUATION PROCESS

Information Useful at Each Step is Indicated in "Ovals"



relating changes to a sequence of possible results. Chapter 3 is devoted to the detailed consideration of this modeling process.

The model that results from this process can then be used for:

(a) Determining whether past research and other "action" projects have filled in gaps in our knowledge of project effectiveness.

(b) Determining what extraneous, uncontrolled changes may exaggerate or conflict with a chosen effectiveness criterion.

(c) Selecting measures of project effectiveness which are reasonable in the light of (a) and (b) above.

(d) Deciding whether a controlled experiment is necessary to measure certain changes.

At this stage of the a priori evaluation process, OHSP has developed a set of project objectives and effectiveness criteria. These can then be compared and integrated with agency objectives which are stated in the proposal.

OHSP is now in a position to evaluate the benefits of the project. This benefit evaluation should consist of two stages:

(a) Evaluation of the value of reaching the chosen measures of effectiveness and evaluation of the technical likelihood of reaching these under ideal implementation.

(b) Evaluation of the agency capability of implementing the project--the personnel, administrative support, and community support.

These separate evaluations should then be combined to obtain an overall benefit measure. Because of the lack of explicit knowledge on relevant causative relations, this measure will in general be non-quantitative. However, for making subsequent cost-effectiveness comparisons it may be desirable to attach a number to this measure of benefit. A hypothetical example will illustrate one way of doing this.

First assign a rating between 1 and 10 to each chosen intermediate measure of effectiveness, where higher ratings reflect the subjective estimates of the importance of the measure. For example, in a driver training project, we might assign 4 to the measure "increase in attitude test score" and 6 to the measure "obtain fewer violations in the first year of driving".

Second, assign fractional numbers between zero and one to each measure reflecting the likelihood of obtaining these intermediate objectives. In the above example, we might assign 7/10 to the first and 1/10 to the second. We then multiply these two numbers and add. In the example we obtain $4 \times 7/10 + 6 \times 1/10 = 34/10 = 3.4$.

Third, assign numbers between 1 and 10 indicating an evaluation of the agency personnel and community support. In the driver education project let us assume this value is 9. We then weight these latter two values by their relative importance. In this example assume each are of equal importance, obtaining a benefit value of $\frac{9 + 3.4}{2} = 6.2$.

The choice of all such weights and ratings is based entirely upon the decision maker's knowledge and subjective opinions. In general we feel that potential project benefit criteria need not be quantified for a priori evaluation. However, if they must for some reason, the above method provides one way of quantitatively integrating these opinions to form one numerical evaluation criterion. It is important to recognize that there are many possible ways to develop such numerical criteria. In addition it is important to use such a criterion with care, for it is a only tool and not a final answer to the evaluation problem.

At this time a cost analysis of the project can be conducted. The big question here is whether the agency can implement its project under the given resources.

Benefits and costs can then be compared to estimate whether the project achieves its benefits in a cost-effective manner. Chapter 6 summarizes the basic theoretical results in this area, and recommends that the classic cost-benefit ratio not be used in project evaluation. Instead a more explicit procedure might consist of subjective comparison of all qualitative potential benefits (or quantitative benefit criteria) with costs. The central questions are whether alternative projects can accomplish these objectives with less resource expenditures or whether such projects can accomplish more extensive objectives for the same expenditures.

These can be answered by making the cost-effectiveness comparisons suggested in Figure 3. These include:

(1) Comparison with alternative solutions for the same problem.

(2) Comparison with competing projects in the same functional area.

(3) Comparison with competing projects in other functional areas.

In this way a ranking of project proposals may be developed. Possibly each person on the staff of OHSP can develop separate rankings. Resolution of the differences in these may yield profitable alternatives and ideas.

Ultimately a list of projects for funding may be developed by selecting projects until funds are exhausted. This selection can be accomplished within those budgetary, geographic and other constraints which may be present. A by-product of the selection will be a set of projects which are returned for re-development or more information. These will then be ranked upon resubmission.

Throughout this section there has been a noticable lack of definite and explicit decision rules and procedures. Instead the emphasis has been what we have called qualitative measures and subjective comparisons. As we have repeatedly indicated, more detailed procedures do not appear reasonable at the present stage of knowledge in this area. Consequently we feel that OHSP should do everything possible to maximize the amount and usage of information available at each step of our "a priori evaluation model". By considering this information in the systematic fashion indicated here, better and more precise decisions can be made in undertaking these very difficult comparisions.

DECISION TO CONDUCT EXPERIMENTAL EVALUATION

There are many projects where it is impossible to make prior estimates of effectiveness without incorporating large amounts of uncertainty. In many of these cases, OHSP or the agency involved should consider the merits of incorporating a controlled experiment into the project to determine these.

An example of such a project is ASMD (Analog Speed Measuring Device). In this project the effect of the device and/or public information on speeding behavior is largely unknown. Hence an experiment to determine this may be very useful.

The decision to conduct an experiment must be made by considering the overall goals of OHSP and the state highway safety program. Experimental evaluation of agency projects should be made only if it is expected that the results of the evaluation will help to guide future projects or to help the agency which is conducting the project. Thus the decision to conduct an experiment must be consistent with the overall goals. In particular the following questions will guide this decision:

(1) What is the value in relation to the overall highway safety program to OHSP and/or the agency of each possible experimental conclusion?

(2) What is the possibility of obtaining experimental control over other factors that might influence the results? -- Can we conduct a good experiment to obtain the desired answers?

(3) What is the cost of this good experiment?

(4) Are the potential answers worth this cost?

Since the development of a good experiment is not an easy task in this area, OHSP and the agency may find it desirable to obtain a consultant's help in answering the above questions--especially question 2.

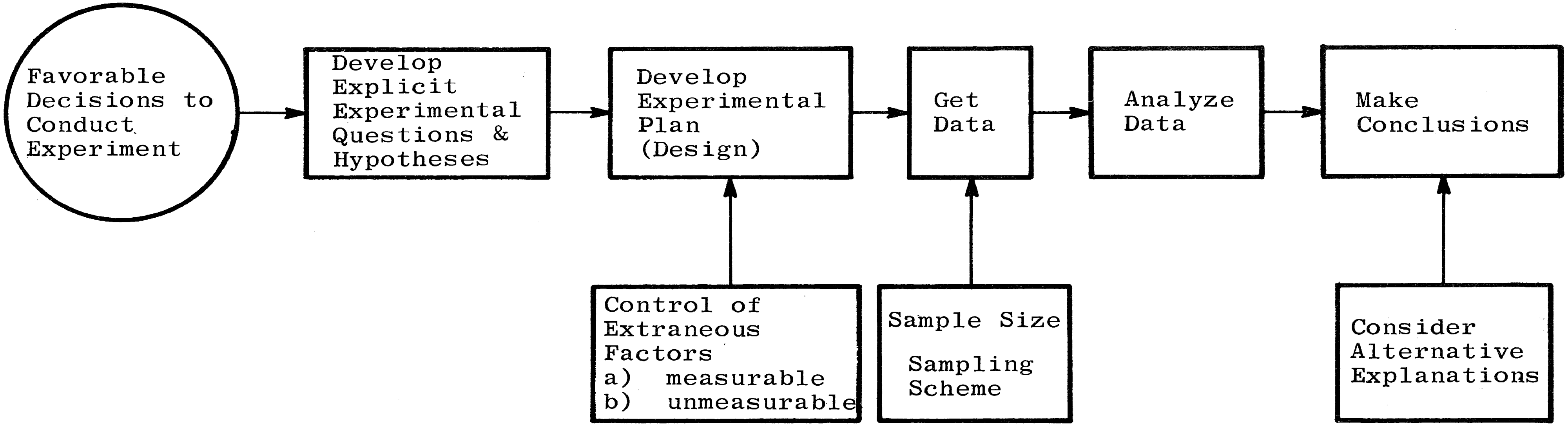
CONDUCT OF EXPERIMENT

If the decision is made to make an experiment part of the project, then an experimental plan must be developed and implemented. Figure 4 shows the relevant steps of this plan.

The design of the experiment is a crucial part of this plan. Chapters 4 & 5 address themselves to this topic. In general, it is very difficult to obtain control over contaminating sources of variation in highway safety projects. In addition, the subsequent experimental data gathering and analysis may involve rather complex mathematical and computational procedures. In such situations a research agency may be able to provide such service on a contractual basis. However, OHSP staff members will profit by understanding the techniques and significance of experimentation in highway safety project evaluation.

Figure 4

CONCEPTUALIZATION OF THE EXPERIMENTAL PROCESS



CONDUCT OF PROJECT

The successful implementation of a project depends on (a) agency capability and interest, (b) agency plan, and (c) agency control.

In general the agency capability and interest will be favorably evaluated before a contract is awarded. Hence we shall not discuss this aspect further here.

Careful and concise project planning in the early stages is essential for several obvious reasons:

(1) To insure that project effects and potential problems in implementation are understood.

(2) To insure that sub-projects take place in an integrated, smooth fashion.

(3) To insure that necessary experimental data is taken (if an experiment is part of the project).

(4) To insure that necessary "status" data is maintained. By "status" data we mean information on the number of persons affected by the projects, general reactions, and general changes being observed.

Because of the importance of such planning we feel that OHSP should monitor the progress of this planning function. This may be done in several ways:

(1) Meeting with agency staff to establish the project plan.

(2) Requiring planning reports from the agency early in the contract period and briefings at regular intervals thereafter.

(3) Encouraging agencies to obtain planning help from any of several highway safety consulting services in the state.

The selection of the specific way to maintain control over the planning process is a policy decision which must be made by OHSP. Various projects may require more extensive planning than others and consequently some blend of the three suggestions above may be desirable.

In order to maintain control over a project the agency must develop some organizational structure to insure that responsibilities for tasks within a project have properly been delegated and maintained. OHSP should evaluate the capability of agency organization to maintain such control. If this is weak, closer guidance from the staff of OHSP is desirable. In any case, it seems reasonable for OHSP to monitor overall progress and provide general guidance by requiring and reviewing written progress reports from the agency at various times. Information to be included in these should be:

(a) Progress during reporting period.

(b) Problems encountered during reporting period.

(c) Financial charges incurred during reporting period.

EX POST FACTO EVALUATION

As a project is completed it is a natural question to ask "what was accomplished?" Such information is useful in avoiding future problems and in planning future programs.

If the early steps of a project are carefully and correctly undertaken then the ex post facto evaluation can be made by comparing the project costs and objectives with those which are actually achieved and observed. Any deficiencies in these should then be examined to determine whether these are due to the nature of the project, the personnel, or other unforeseen factors. If possible the staff at OHSP should consider solutions which would have alleviated such deficiencies had they been implemented. In this way similar problems will be avoided in the future.

On the other hand, observed benefits of the project which were not anticipated in the early statement of objectives should be recorded for future planning. Notice however, that the project cannot be evaluated ex post facto by determining its success in achieving objectives which were first considered after the project planning was completed. For instance, if a project objective is stated as "installing computerized control to reduce the cost of road signal maintenance", and if a project is planned to implement this, then the project should not be evaluated by looking at "changes in traffic flow". If a favorable change in this latter measure occurs, it should be viewed as an extra benefit. This again points out the importance of clear prior formulation, development, and planning.

With this overview in mind, we now turn to a detailed discussion of the technical material the staff of OHSP should be familiar with to implement the concepts of this section.

2. APPLICATION OF SYSTEMS ANALYSIS TO PROJECT CLASSIFICATION AND EVALUATION

In this chapter we shall expand upon the early steps in our a priori evaluation plan (Figure 3). Specifically we shall discuss:

(1) The project classification plan.

(2) The concepts underlying a sub-system model.

(3) The use of this sub-system model in delineating potential project effects and in formulating evaluation criteria.

In order to do this we first must develop a brief systems conceptualization of the highway safety problem. This conceptualization points out the need for our systematic evaluation plan and serves to guide project analysis.

SYSTEM CONCEPTUALIZATION

The problems of highway safety improvement are far more complex than many early investigators recognized. This complexity is due to at least three major factors:

(1) The large number of variables influencing highway safety.

(2) The complex and in most cases unknown relationships among these variables and between these variables and "safe performance".

(3) The fact that both these variables and functions vary probabilistically over time.

Because of this complexity, it has become necessary to structure highway safety research and program evaluation within the context of a "systems analysis" framework. Such a framework provides:

(1) A basis for breaking down the complex system into components which are more amenable to scientific research.

(2) A basis for investigating the interrelationships that exist between these components.

(3) A basis for allocating research effort to the highway safety problem.

In this section we shall explore this systems conceptualization as it pertains to program evaluation. Specifically we have developed a three-way classification plan for projects. This plan highlights the major system components and project objectives involved, and serves to guide the subsequent evaluation plan. The evaluation plan logically organizes the steps of a project analysis.

MULTIVARIATE STRUCTURE OF PROBLEM

As we study an individual collision, we can identify many factors surrounding the event, "an accident". Many of these fall into the group of potential causal factors. As H. H. Jacobs has stated (7):

> "Remedies which would have prevented an individual accident are almost always apparent from even casual investigation".

The apparent ease of identifying potential causal factors has caused many persons to develop their own "optimal solution" to the problem of accident reduction. This apparent ease is also a major shortcoming. Many potential solutions have been proposed but there is not a solid base of evidence to support a set of "best" solutions. We feel that there are many variables which contribute to accident occurrence. In fact, it is reasonable at this point to assume that most of the factors indicated as causal factors do contribute to the creation of a potential collision situation. A crash occurs whenever the combined contribution of these factors exceeds some critical level. Thus, it is the combination of factors that finally cause the event--a crash. The research to date has not revealed a single factor or a unique combination of factors which are involved in most of the accidents. Several carefully designed studies have indicated the overinvolvement of certain factors. (e.g., alcohol by Borkenstein (3), et al.) However, much research remains to be done in this area.

The above discussion points up the danger in attempting to evaluate the effect of modifying a single factor which contributes to the event--a crash. In particular, the Federal program is designed to deal with specific factors. Therefore, developing a project evaluation plan involves the above problems. This does not mean that we should wait for more knowledge. However, it does mean that an evaluation plan must respond to these problems.

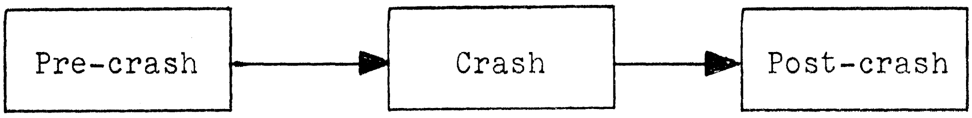
In order to measure the effect of a particular project, there must be knowledge concerning the influence of other factors which are not being modified. Ideally, we would like to fix all other factors and thus eliminate their effect so that the proposed projects can be evaluated directly. In some cases, (e.g., modifying a dangerous intersection) it may be possible to control enough factors so that a direct measure of loss due to system breakdown can be made. This direct measure is, of course, the best approach since it is the ultimate objective. However, in most situations (e.g., driver training, law enforcement, etc.), it will only be possible to evaluate some intermediate objective. The selection of this intermediate objective is related to the particular problem. For example, the evaluation of a driver training project might measure one of the following:

1. The amount of knowledge transferred by means of a written test.
2. The effect of this knowledge on a controlled driving situation.
3. The change of attitude as measured by a psychological test.
4. The driving records of persons who have taken the course.
5. The effect on number of collisions for all persons in the community in the age group given driver training.

The early items on the list are more directly related to the specific program and consequently provide measures of project effectiveness which are less affected by intervening variables. However, the question of how this measured effect relates to the final objective is less apparent. If item five is measured accurately, we do have knowledge on the ultimate effect. Still, the number of uncontrolled variables operating at that level will reduce considerably the accuracy of interpreting the true cause and effect relationship. An evaluation of the advantages and shortcoming at each evaluation level must be performed. Subjective knowledge will, of course, influence the final decision.

CLASSIFICATION PLAN

The traffic collision process may usefully be split into three time-oriented phases as shown below:



The pre-crash phase includes those factors leading up to and influencing the occurrence of the crash. The crash phase includes the response of the vehicle/operator to the collision. The post-crash phase includes recovery, clean-up and related procedures. Specific events in these phases are interrelated. For instance high speed in the pre-crash phase generally causes a much more violent vehicle/ operator response in the crash phase. Although it is useful to consider the phases separately, their joint effects must also be considered.

The overall objective of highway safety effort is to reduce the loss due to system breakdown. The accomplishment of this objective might be initiated by allocating all effort to one of the phases indicated above. For instance, one might conclude that collisions are unavoidable. Then all effort could be spent making super-strength vehicles and padding the driver to avoid injury. Most researchers today argue that such an allocation would be sub-optimal. Instead the common feeling is that effort should be allocated to all three phases, although the relative efforts to be extended on each phase are still a subject of debate.

Thus the overall objective of highway safety effort involves three intermediate objectives. In the pre-crash phase we wish to reduce the "potential" for system breakdown. Given that the system does break down, in the crash phase we wish to reduce the severity of this breakdown. Finally in the post-crash phase we wish to reduce the time until the system returns to normal and upgrade effectiveness of care in this time interval.

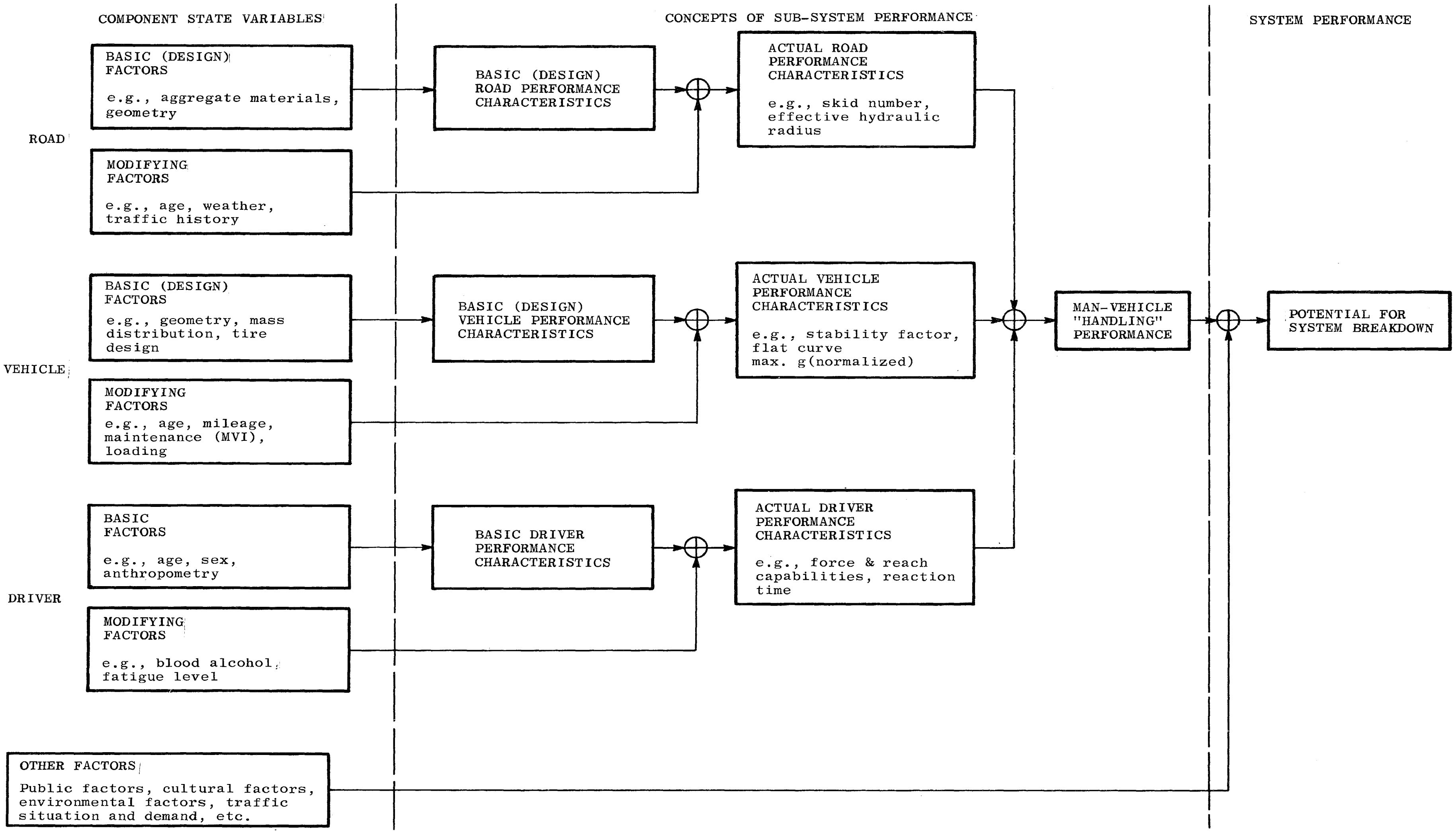
Within each phase, the system may be subdivided into three interactive components: the vehicle, the operator, and the environment. This subdivision is outlined for the pre-crash phase in Figure 5.* The model relates potential for system breakdown to the three components, which interact through performance and control characteristics. Similar models can be evolved for the crash and post-crash phases.

---

*Based on a conceptualization by Howard Dugoff of Highway Safety Research Institute, The University of Michigan, Ann Arbor, Michigan.

Figure 5

PRE-CRASH SAFETY: A SIMPLIFIED SYSTEMS MODEL



The important point to be obtained from this componentwise subdivision is the complex, interactive relationship between system variables and the intermediate within-phase objectives. Consequently the effect of changes in system variables on system performance may not be directly measurable. For instance, in the pre-crash phase, the effect of a new driving simulator on reducing potential for system breakdown may be masked by the effects of differences in operator socioeconomic characteristics, vehicle design characteristics, and other extraneous changes in the system. Therefore the choice of intermediate levels of effectiveness and careful control of intervening variables are essential for accurate project evaluation.

In addition to classifying projects by system phase and system components, it is useful to incorporate a third categorization--the action resulting from a project. In this section we shall outline our approach to this classification. Specifically we suggest that highway safety projects involve three types of action: informational, direct component change, and indirect component change.

INFORMATIONAL PROJECTS. Projects in this category provide increased knowledge on the highway safety system by obtaining data from the system and by presenting it in useable form. Examples of such projects include the Department of State driver record system and the State Police accident report system.

Informational projects are designed to provide information describing system parameters and operation. Included in this category are "state of the art" surveys, information retrieval projects, and information seeking projects (e.g., engineering studies for planning purposes.) The information from these is utilized for decision-making purposes about the operational aspects of the system. Two basic uses of such information are (a) developing an overall system description (number of miles traveled, number of collisions, number of deaths, etc.) and (b) identifying specific problem areas (high accident roadways, high violation operators, etc.) Under (b) two specific types of information may be made routinely available by a project. One type identifies those problem areas that are "out of control." For instance one may need to know which intersections have the largest number of collisions. A second type provides rapid access to information on specific parameters. For instance one may need to know the collision rates on a series of specified intersections.

The specific objectives of informational projects generally will be framed in terms of the quality and quantity of information obtained. Quality may be further sub-divided by considering accuracy and retrieval speed.

COMPONENT CHANGE PROJECTS. The second action categorization includes those projects which propose to make some change on the system. Generally the investigator hopes that this change will improve system performance. Within this category there are two major project groupings -- direct component change and indirect component change. These two types of projects affect the system differently and consequently must be evaluated differently.

Figure 6 indicates the general relationships between these project categories and the ultimate pre-crash objective--reduction in potential for system breakdown. The development of the chain of events relating a component change to system objective is basic to the evaluation of projects involving project change. We shall denote this chain of events the "causal chain." The "length" of this chain depends on the number of intermediate changes that take place as a result of project implementation and on the number of intervening, uncontrolled variables affecting the sub-system of interest. The length of the chain and the degree of knowledge relating a system change to system objective determine whether a project involves a direct or indirect component change.

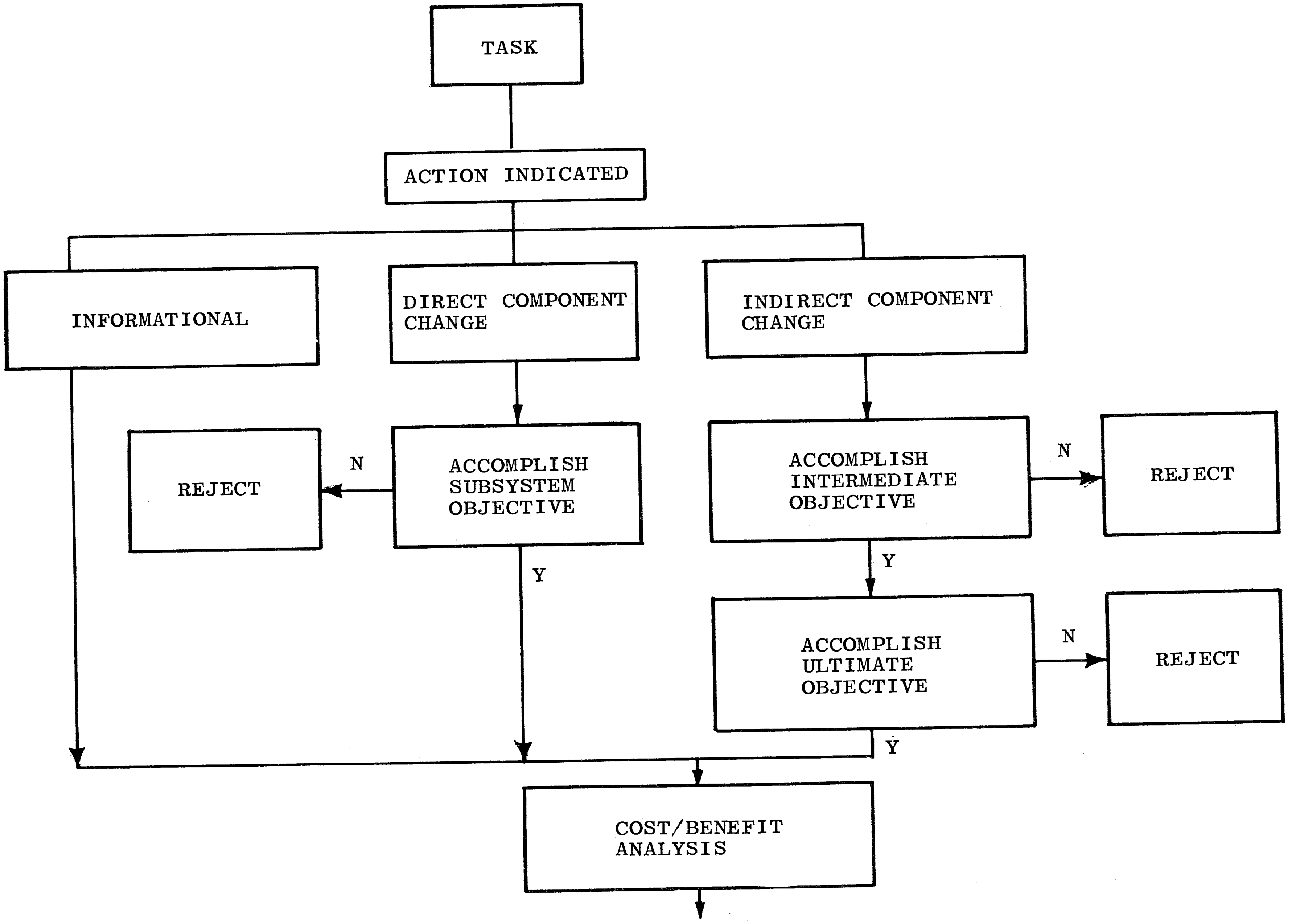
Direct Component Change Projects. Projects in this category can be expected to have a direct effect on system performance if they successfully accomplish the proposed change. For example, knowledge on vehicle dynamics and operator handling capability may indicate an upper limit on safe turning speed at a particular sharp turn in the roadway. Collision experience at this turn indicates an "overrepresentation" of crashes involving loss of control. Suppose, a specific project proposes to modify the road (the environmental component) in order to correct this problem. Indeed, several alternative modifications might be relevant. If this project is carried out so that controlability is measurably improved, we anticipate with high "probabillity" that the frequency of this class of crashes will be reduced.

Direct Component Change projects take advantage of prior knowledge and a shorter "causal" chain. Consequently the change in system performance can be inferred from the known relationships within the chain. The evaluation of such projects may be made using overall system performance as a direct, measurable objective. This does not imply that additional system variables can be ignored in the project evaluation, for their effects, if left uncontrolled, may mask the system measure of performance.

Returning to the example on roadway modification, suppose the investigator decides to add warning signs to make operators anticipate the turn. At the same time the population of operators passing through the turn changes, with the new drivers having more careful driving habits. Then the subsequent reduction in crashes may be due to: (a) the signs, (b) the more careful operators, or (c) both. If, on the other hand, the population changes so that drivers have less perception and awareness, the collision rate may go up -- not because of the signs but because of the operators. In both cases failure to consider the degree of operator change leads to inconclusive or erroneous results in the evaluation.

Indirect Component Change Projects. Projects in this category propose system changes whose effects are related to the system objective through a long, complicated and loosely defined chain of events. Thus changes in intermediate levels of performance (specific project objectives) -- even if successful -- do not necessarily relate to the overall system objective. Figure 6 points out this problem. For example, consider a specific driver education project which proposes to add driver simulators to a curriculum. The immediate objective is increased knowledge and skill in the student. The ultimate objective is a reduction in loss due to system breakdown which can be attributed to this change. Thus, by considering the causal chain, the project might be evaluated in terms of the increase in student knowledge and skill as measured over time by various testing devices. Careful methodology is essential for attributing changes in these responses to changes in the curriculum. However, even though these changes are found to be "real" we are unsure that they will directly affect (reduce) potential for systems breakdown. This is true because of the large number and uncertain effect of intervening events between an increase in knowledge and this objective.

Figure 6. THE "CAUSAL CHAIN" BETWEEN PROJECT TASK AND OBJECTIVE



In summary, we have developed a classification plan for projects utilizing the concepts of systems analysis. This plan places projects into three categories: phase of the crash process, component of the system effected, and action dictated by the project. The essentials of the plan are shown in Figure 7. Within each of the three categories projects may be classified into one of three levels, resulting in 27(3x3x3) possible classifications.

The classification plan serves to guide the initial steps of an evaluation by objectively reducing the system to meaningful subsystems and showing in gross terms the intervening and extraneous variables which must be controlled for meaningful evaluation. The "action" classification introduces the idea of a "causal" chain of events and considers the problem of "length" within this chain. These concepts lead to consideration of a general evaluation plan, which must be dependent upon this project classification in order to meaningfully respond to these problems.

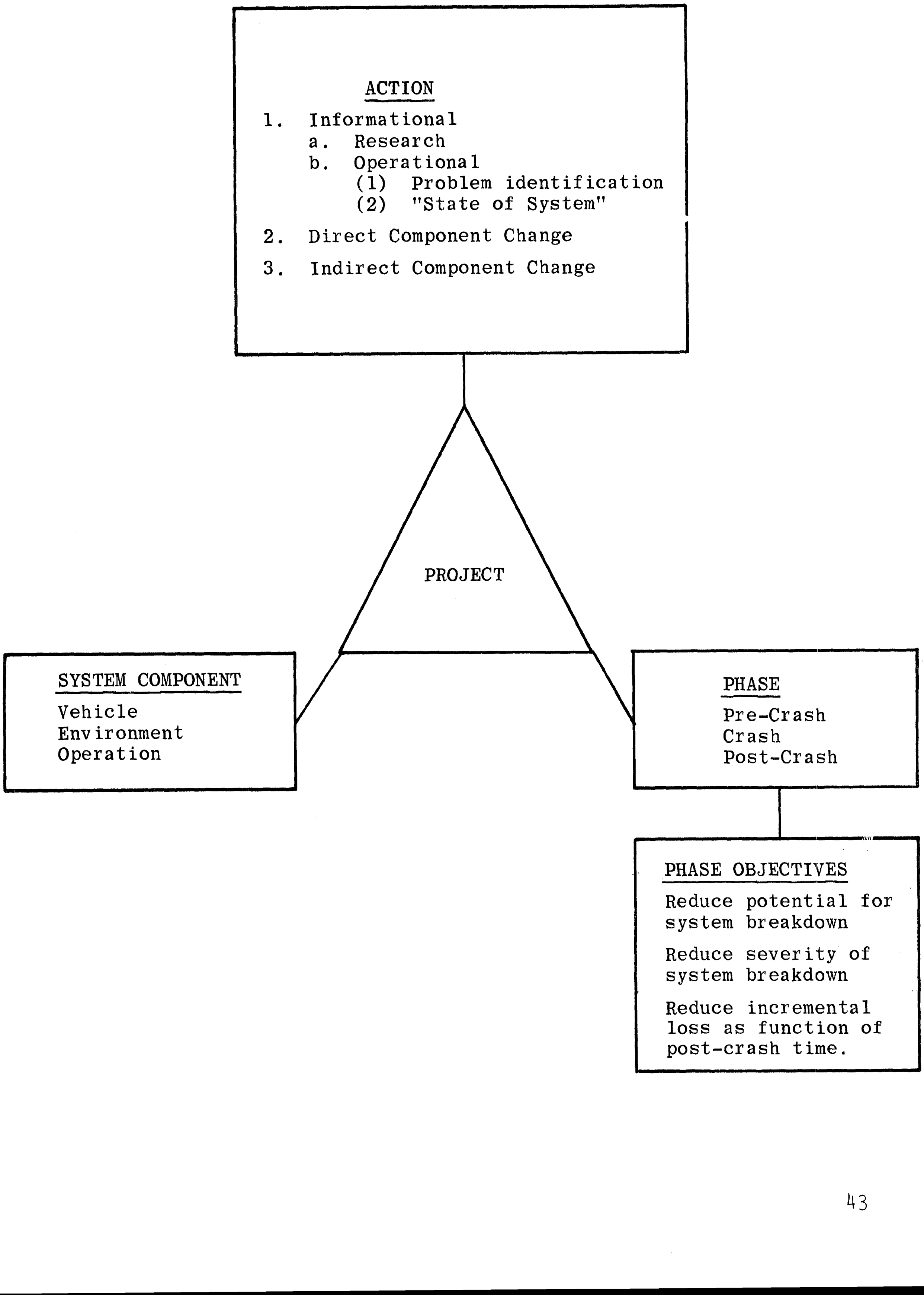
CONCEPTS UNDERLYING A SUB-SYSTEM MODEL

As we indicated in the introduction, the objectives of project evaluation are three-fold:

(a) To insure that the internal structure and design of projects is such that meaningful conclusions can be obtained if the investigator carries out his plan as stated.

(b) To make comparisions between projects in related functional areas within a valid cost-effectiveness framework.

Figure 7. THE CLASSIFICATION PLAN



(c) To make comparisons and allocate funds to functional areas within a valid cost-effectiveness framework.

In Chapter 1 we outlined a general a priori evaluation plan which is directed at these objectives. Figure 3 outlines the relevant steps in this plan. Our objective in this section is to expand upon steps two and three of this plan--the development and use of a sub-system model.

Step two in the plan involves the construction of a sub-system "model" delineating the pertinent variables and relationships which relate to the phase objective. The classification plan will aid in identifying the major factors involved. The appropriate sub-system for modeling may be a functional area or some defined subset of a functional area. For instance, for a driver education project we may develop a conceptualization of the functional area "driver education". On the other hand, for a highway improvement project we may develop a model only for the specific physical environment involved.

The development of such a model requires extensive use of prior "state of the art" knowledge, the construction of hypothesized relations, and careful consideration of all potential variables. Consequently, the process is iterative, for at any time our knowledge of these complex processes is incomplete. Hopefully, research work will continue to fill in these gaps. Careful modeling may in fact reveal research deficiences.

The specific considerations and techniques which are of use in abstracting and constructing this sub-system model are discussed in Chapter 3. At this time it is necessary only to recognize the importance and priority of undertaking this step.

USE OF THE SUB-SYSTEM MODEL

After a model conceptualization of the relevant sub-system has been developed we proceed to examine the potential effect of the project of interest on this sub-system (step three). We first determine where in the causal chain the project acts. Then we determine how the project acts: which variables and relationships are affected and how these are affected. The quantitative aspects of this second question are discussed in Chapter 3.

We then examine factors in the causal chain which are "dependent" or related to those manipulated by the project. In most cases there is a hierarchy of such events which are logically ordered. It is important to observe that changes resulting from a project should result in changes in each "event" following the project in the causal chain. When these intermediate changes contribute to the system objective, we accomplish "sub-objectives" of the project. Because of the logical relations, all sub-objectives must be met before the sub-system objective can be met.

At this stage we must select some step in the causal chain upon which to build our evaluation. The determination of this level is a factor of:

(1) The number and hypothesized effects of uncontrolled, intervening variables and relations. If we can adopt a research strategy to control or eliminate some of these our evaluation level may generally be closer to the ultimate subsystem objective.

(2) The ability to measure response at each level.

(3) The ability to generate a sufficient sample for meaningful evaluation at each level.

(4) The length and degree of knowledge relating project objectives to the system objective.

Within the chosen level of response we then select particular variables which will serve as evaluation criteria. The project will then be internally evaluated in terms of the change in these criteria. Subsequent comparisons will be made by comparing the magnitude of this change for various projects in the same functional area. For comparisons between functional areas different variables must be equated into common terms. This latter problem will be discussed later in this report.

3. DEVELOPING A SUB-SYSTEM MODEL FOR PROJECT ANALYSIS

In this chapter we shall examine step two of the a priori evaluation plan in greater detail, for the development of an adequate sub-system conceptualization is the key to successful project evaluation. Specifically we consider techniques for abstracting the essential sub-system components and parameters in a way so that project effects on the subsystem can be represented in a logical sequence. Analysis and experimentation can then be performed using the relationships defined by this sequence. In this section we present the general basis of this modeling process and indicate alternative levels of detail which can be incorporated into such a model. This detail depends upon the investigator's prior knowledge, the level of detail necessary for making adequate decisions, and the economic and technical ability to apply increasingly sophisticated modeling tools.

Although we present a survey of concepts useful in model building, we recognize that this section gives only nominal guidance on "how" to model specific sub-systems. This latter question can best be answered by developing sample models, and we propose to answer it in this way in the application section of our work.

GENERAL CONSIDERATIONS IN THE MODELING PROCESS

The general concepts in developing a model of some system are outlined in Figure 8.

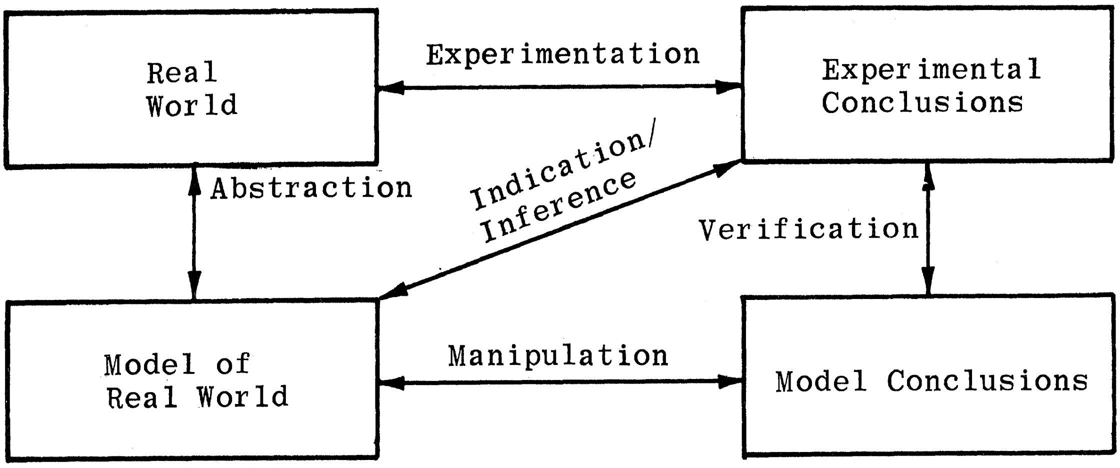


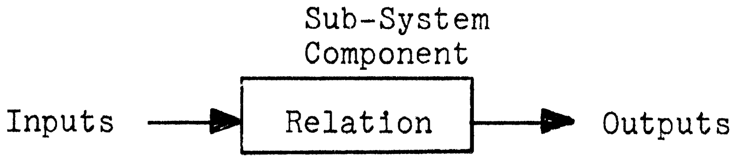
Figure 8. CONCEPTS IN DEVELOPING A SYSTEMS MODEL

The basic goal in abstracting a real process is to obtain a simplified representation of the process which is realistic, general, economical, and tractable. In particular, we desire a representation with the smallest number of parameters which still satisfies these four goals. Observe that these goals are incompatible. For instance increased realism must be sacrificed for increased generality. Similarly economic considerations may dictate decreased attainment in the three other goals. The necessary compromises in this modeling process should be based on the projected use of the model, the ability to estimate and manipulate the chosen parameters, and the "state of the art" knowledge on the relevant sub-system. These trade-offs should be considered early enough in the model building process to prevent over-zealous pursuit of any subset of the objectives.

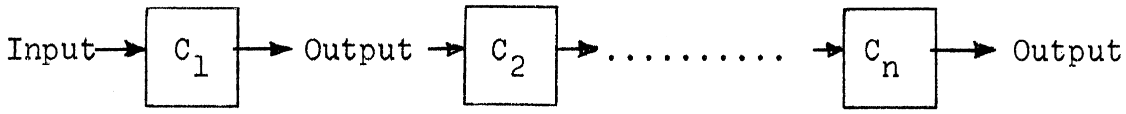
COMPONENTS OF THE SUB-SYSTEM MODEL

We return now to the specific problem of developing a sub-system model useful in highway safety project evaluation. In this section we shall outline in general the components of this model. Recall that we have conceptualized the subsystem in terms of a logically ordered sequence of components, where the ordering is in terms of the ultimate phase objective. For instance, consider the driver education functional area. Here we have a course transmitting information to a student who may respond through an immediate internal change. This in turn may induce external changes which may induce subsequent changes in the pre-crash phase goal -- the potential for systems breakdown. Each of these major components can be conceived as a "black box", accepting inputs, transforming them, and yielding outputs. In the case of the course and student, these "black boxes" have physical representations. However in the case of internal and external change and potential for systems breakdown, no such physical analogs exist.

A component may be represented as in the diagram below:



For instance, consider the "course" component of the driver education functional area. Inputs to this include the instructional material, the instructor, and teaching devices, while outputs include course structure, data, and presentation. In general the inputs to a component are called dependent or response variables. Of course these dependent variables may be the independent variables for the next component in the causal chain. This chain may then be represented as shown in the following diagram:



Each sub-system component transforms the input variables into output variables. Thus, we might conceive of a component as a set of functional relationships* expressing the relationship between certain input independent variables $x_1,\ldots,x_r$, and output dependent variables $y_1,\ldots,y_s$. In mathmatical notation these functional relationships might be expressed as

$$y_i = f_i\ (x_1,\ldots,x_r) \qquad i = 1,\ldots,s \tag{1}$$

That is the ith output variable y_i is a function of r input variables $x_1,\ldots,x_r$. Here r and s are a pair of integers indicating the number of inputs and outputs respectively.

Returning again to the "course" components of driver education, we may write

$$\begin{bmatrix} \text{course presentation} \\ \text{course structure} \\ \text{course information} \end{bmatrix} = \text{function of} \quad \begin{bmatrix} \text{instructional material} \\ \text{teaching devices} \\ \text{instructor} \end{bmatrix}$$

Our ability to identify and exactly determine the set of functional relationships associated with a particular sub-system component depends again on the level of prior knowledge available. In many cases we can do nothing more than indicate our best "guesses" as to the input and output variables. With more knowledge we may be able to make statements about the direction of change in a particular y_i that comes

---

*The reader should take care to distinguish between "functional project area", as implied by the State Manuals and "functional relationship", a mathematical concept.

from a change in certain of the $x_1,\ldots,x_r$. In terms of the driver education example, we may be able to infer that a simulator increases course information output or information clarity. More sophisticated knowledge may allow one to make statements about the magnitude of such a change as well as the direction.

At each stage of the iterative model building process, the functional relations (1) may be written in more detail as

$$y_i = f_i\ (x_1,\ldots,x_m;\ x_{m+1},\ldots,x_r) + E \qquad i = 1,\ldots,s$$

where

y_i = ith output, dependent variable from the subsystem component.

f_i = ith functional relation relating sub-system inputs to sub-system outputs.

$x_1\ldots,x_m$ = those input variables which are specified and are capable of measurement.

$x_{m+1},\ldots,x_r$ = those input variables which are not specified in the model or are not capable of measurement in the process.

E = the residual error not explained by the functional relation. This may be due to (a) measurement error and/or (b) inadequacy of the functional relationship in describing the true response of y_i to the input variables.

We shall now proceed to examine the concepts of variables, functional relations, and residual error in more detail.

VARIABLES. We have seen that a subsystem can be described by a series of input-output relations, where within each component input variables are related to output variables. These variables can be measured in different ways. These include nominal, ordinal, and interval measurements defined as follows:

Nominal - A categorization into groups that are different in some way. For example, sex is a nominal variable taking on two qualitative values, male or female. Such a variable may be quantified by letting the variable take the value 1 if male and 0 if female.

Ordinal - A categorization into groups that differ in an ordered way. For example, educational level is a ordinal variable taking on values for elementary school, secondary school, and college. Here we may quantify the variable in an ordinal way by letting 1 = elementary, 2 = secondary, 3 = college.

Interval - An ordering where the relative values assigned to each group indicate the degree of difference. For example, consider the variable "speed." A vehicle traveling at 60 miles per hour is moving twice as fast as a vehicle traveling at 30 miles per hour.

Interval variables may be further sub-divided into discrete variables and continuous variables. Continous variables exist at any point over some range. For example, temperature and pressure are continuous interval variables. Discrete variables occur only at pre-assigned values over some range. For example "number of crashes is a discrete interval variable taking on the values 0, 1, 2, ..., N, where N represents an upper limit.

The way in which a variable is measured depends both upon the conciseness of definition and upon the measurement technique. Consider the dependent variable "potential for system breakdown". We might define this variable as the "probability" that a specific man-vehicle-environment will have a crash in the next year. Then the variable is a continous interval variable over the range zero to one. Alternatively we might define the variable as the number of crashes in a year: then we have a discrete interval variable.

In studies involving human populations, we are often unable to precisely define the variable of interest. For instance consider the variables attitude, behavior, and socio-economic status. We also have limited measurement capability in determining these once they are defined. Consequently we are generally forced to deal with nominal or at best ordinal measurements.

FUNCTIONAL RELATIONSHIPS. We turn now to a discussion of the functional relationships relating inputs to outputs. These relationships serve as convenient and useful ways of abstracting knowledge of the "real world." However, for many sub-systems of interest our knowledge is so limited that postulating exact relationships is impossible. Instead we must be content with some knowledge of the direction and magnitude of change in an output due to change in some subset of the inputs.

Because of this the utilization of functional relationships in modeling "cause-effect" phenomena must be based on their predictive power rather than their closeness to reality. That is, we must base our selection of a functional relation on the size of the deviation of the actual response from the predicted response. As research in highway safety progresses we hope to be able to increase our modeling sophistication -- moving from predictive models toward models which more accurately represent the underlying causative relations.

Within this set of predictive relations the choice of a specific functional relationship to model some phenomena can still be made at one of several levels of detail. When little prior knowledge is available, empirical data analysis techniques may be utilized to search out an adequate predictive relationship. The available set of tools for doing this is rapidly expanding and their potential for highway safety research is just beginning to be uncovered. In using these tools one must keep in mind that any set of data can be fitted exactly with some model. For instance, consider the "scatter plot" illustrated in Figure 9. Two functional relations $f_1(x_i)$ and $f_2(x_i)$ are illustrated which "fit" the data exactly. Notice, that if another data point x*

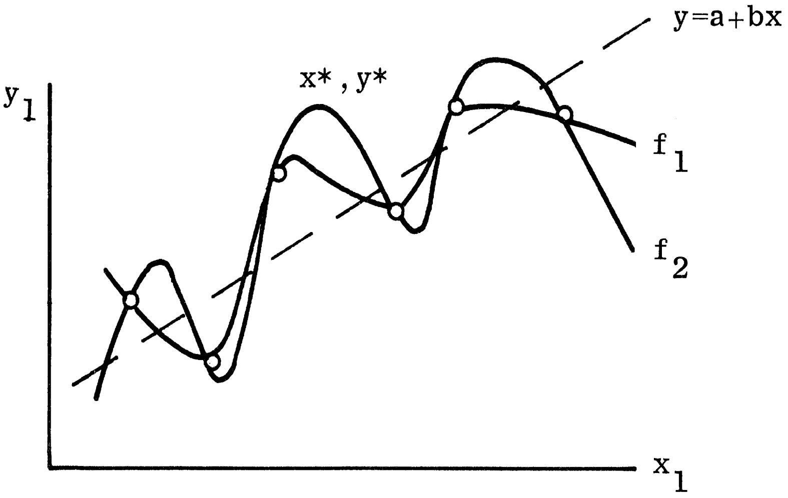


Figure 9. THE PROBLEM OF FITTING FUNCTIONS TO DATA

is taken, neither model predicts y*, the response to this value.

We are interested in selecting a functional relationship that represents the response of interest adequately with a minimal number of parameters. Using the terminology of Dr. John Tukey, we are interested in "parsimony" in our selection of a model. The functional relations f_1 and f_2 above are higher degree polynomials in the independent variable x; thus they have several parameters which must be estimated. On the other hand a linear relation $y_1 = a + bx_1$ has only two parameters (a and b) and seems to predict all values with reasonable accuracy.

Most of the functional models used in experimental design have this linear structure. The simple linear model in one independent variable can easily be extended to a model in m variables:

$$y = a_1x_1 + a_2x_2 + \dots + a_mx_m$$

Here m parameters $(a_1, \dots, a_m)$ must be estimated in order to use the model.

The use of a linear model may be acceptable because it adequately "approximates" the true response. Furthermore certain nonlinear functional relations can be transformed into linear relations, and complex relations may be approximately linear for small changes in the independent variables. Consequently, the selection of a linear relation may serve our purposes adequately, even though we recognize its limitations and approximate nature. In terms of the goals of abstraction mentioned earlier, we may sacrifice reality and generality for increased tractability in visualizing and estimating and for increased economy in development and utilization.

Returning to the general situation, we have seen that the choice of a functional relationship involves the deletion of certain variables and the approximation of certain relations. Then, in the general case we may write

$$y_i = g_i(x_1,\dots,x_m) + h_i(x_1,\dots,x_m) + k_i\ (x_{m+1},\dots,x_r)$$

where

$g_i(x_1,\dots,x_m)$ is the assumed relationship for these variables which are specifically included and measured. In the case of a linear model $g_i(x_1,..,x_m) = a_1x_1 + \dots + a_mx_m$.

$h_i(x_1,\dots,x_m)$ is the secondary, unmeasured relationship for the measured variables $x_1,\dots,x_m$. For instance, if the "true" relationship among the variables is quadratic (proportional to x_2), then $h_i(x_1,\dots,x_m) = b_1x_1^2 + b_2x_2^2 + \dots + b_mx_m^2$ where $(b_1,\dots,b_m)$ are the unknown coefficients)

$K_i(x_{m+1},\dots,x_r)$ is the unmeasured relationship for those variables that are not included in the model.

In reality $h_i(x_1,\dots,x_m)$ and $k_i(m_{m+1},\dots,x_r)$ are lumped together as "equation error". One objective of experimental design is to keep such equation errors within the bounds necessary to make valid predictions from the assumed relationship.

ERROR. In addition to the "equation error" in specifying a functional relationship, measurement error also may distort the independent and dependent variables. Measurement error is directly related to the precision of a variable's definition and the degree to which this definition is observed in gathering data. In typical physical systems (a chemical reaction process, for instance) variables are in general precisely defined and measurement error is almost wholly dependent on instrumentation accuracy. As we have seen, human variables suffer from both definitional and observational problems. Specifically, we are unable to accurately and quantitatively define variables such as attitude, intelligence, and behavior. Therefore we select either nominal or ordinal variables. This approximate nature introduces certain measurement error. In addition, response error and bias may be undetected by the measuring device (for instance, a written test). Furthermore, inaccuracy in recording responses may be present, as in the case of interviewing.

Thus several sources of error may combine to give an inaccurate "reading" of the variable:

$$\begin{bmatrix}\text{measured}\\\text{variable}\end{bmatrix} = \begin{bmatrix}\text{true}\\\text{variable}\end{bmatrix} \pm \underbrace{\begin{bmatrix}\text{definition}\\\text{error}\end{bmatrix} \pm \begin{bmatrix}\text{response}\\\text{error/bias}\end{bmatrix} \pm \begin{bmatrix}\text{recording}\\\text{error/bias}\end{bmatrix}}_{\text{variable error}}$$

In many cases we lump these error sources together as variable error. A functional relation can then be written as:

$$\begin{bmatrix}\text{true response}\\\text{variable}\end{bmatrix} \pm \begin{bmatrix}\text{response}\\\text{variable error}\end{bmatrix} = \begin{matrix}\text{assumed}\\\text{function}\end{matrix} \begin{bmatrix}\text{independent}\\\text{variables}\end{bmatrix} \pm \begin{bmatrix}\text{independent}\\\text{variable errors}\end{bmatrix} \pm \begin{bmatrix}\text{equation}\\\text{errors}\end{bmatrix}$$

In functional notation

$$y_i \pm Ey_i = g_i(x_1 \pm Ex_1, \ldots, x_m \pm Ex_m) \pm h_i(x_1, \ldots, x_m) \pm K_i(x_{m+1}, \ldots, x_r)$$

We may further lump response variable errors, independent variable errors, and equation errors together obtaining

$$\begin{bmatrix}\text{true response}\\\text{variable}\end{bmatrix} = \begin{bmatrix}\text{assumed}\\\text{function of}\end{bmatrix} \begin{bmatrix}\text{independent}\\\text{variables}\end{bmatrix} \pm \begin{bmatrix}\text{combined}\\\text{error}\end{bmatrix}$$

or in functional notation $y_i = g_i(x_1, \ldots, x_m) + E$

The techniques of experimental design and data analysis are designed to control or compensate for these combined sources of error in order to obtain meaningful conclusions. These techniques will be explained further in Chapters 4 and 5.

4. OVERVIEW OF EXPERIMENTAL DESIGN

In many projects the large gaps of knowledge in understanding and accurately measuring countermeasure effectiveness suggest that an experimental evaluation should be part of the project. As we have indicated in Chapter 1, the decision to conduct such an experiment can be made by answering four questions:

(1) What is the value to the overall highway safety program of OHSP and/or the agency of each possible experimental conclusion?

(2) What is the possibility of obtaining experimental control over other factors that might influence the results?

(3) What is the cost of a good experiment?

(4) Are the potential answers worth this cost?

If an experimental evaluation is decided upon, a valid experimental plan (design) should be developed before the project is initiated. This development is a difficult task, for it requires consideration of potential project effects, the effects of uncontrolled factors, and the problems induced by using a "sample" for evaluation. Our objective in this section is to provide an introductory overview of this process.

The objective of experimental design is to obtain useful information with minimal effort. As such, experimental design is closely related to the scientific method -- proposing a hypothesis and the evaluation of this hypothesis by means of experimentation. Defining and structuring project objectives -- using the concepts of Chapters 2 and 3 -- are necessary pre-conditions to experimental design. These lead to the selection of particular observational and analytic techniques before experiments are conducted. Collecting data with the "hope" of doing something usually limits the potential usefulness of an experiment.

We have seen that experimentation is useful in verifying models abstracted from the real world and in estimating the parameters of such models. While we shall not explore the technical tools for performing these, we shall consider problems in designing an experiment to obtain the relevant data. In this section we shall briefly explore the classical components of design -- replication, randomization, local control, and estimation of interaction -- as they relate to traffic safety. In Chapter 4 we shall detail other specific techniques and tools of particular usefulness in traffic safety.

REPLICATION

Replication is synonomous with repetition. By "holding" independent variables at fixed pre-selected values and repeatedly measuring the response variables, one is able to obtain some estimate of the error term associated with the relationship under consideration. For instance, consider simple model predicting collision rate at an intersection (y) as a function of time of day (x_1) and traffic density (x_2). The model is then

$$y = f(x_1,x_2) + E$$

where f is some functional relationship (for instance, a linear relation $a_1x_1 + a_2x_2$) and E is the combined equation and measurement error.

If we could fix the time of day and traffic density and measure collision rate for some period of time, the variation in this rate would be an estimate of the error E.

Unfortunately replicated experiments are often impossible in highway safety work, for many times it is impossible to fix independent variables at pre-specified levels. In context of the above example, we may fix the time of day by observing the accident process only within specified time intervals. However, traffic density within these time intervals will vary from day to day. Hence we cannot exactly repeat the experiment. Instead, we must consider the joint variation of collision rate and traffic density. This involves the use of techniques of multivariate statistical analysis and in general complicates the problem.

RANDOMIZATION

Throughout this discussion we have stressed the complex, multivariate nature of the highway safety problem. In a particular study or evaluation there is some set of variables which the investigator is explicitly considering and a much larger set of variables which he is not considering. The effects of this set will bias or decrease the precision of the desired measure of project effectiveness.

One method of increasing "control" over this set without specifically measuring it is to use randomization in assigning project levels to the "experimental units." Randomization is a process which insures that all measured variables have an "equal likelihood" of being affected by unmeasured causes. For example, consider a simulator in driver education. This project has two levels - no simulator or simulator. In this case the appropriate experimental unit is a student. We wish to utilize simulators in training some students and at the same time train other students without this device. A systematic selection of those students to receive the simulator training may bias the results. For instance, suppose that boys are given the simulator and girls are not. Then the response (evaluation criteria) reflects both a "simulator effect" and a "sex effect" if they exist. If, on the other hand, the simulator was "randomly" applied to a population of students, the bias due to extraneous factors tends to average out when the simulator effect is estimated from the data.

A second use of randomization arises in assigning interviewers to populations. For instance, if data concerning a group of alcoholic and a group of non-alcoholic operators were being collected, the assignment of one person to each of the populations might in fact introduce an interviewer bias into the results. On the other hand, assigning members of the two populations to the interviewers on a random basis eliminates this source of error.

Randomization may also be used to control the effect of variables which can be defined but which are not measured due to excessive cost. One example of this type of variable arises in classifying drivers. (9) It seems that knowledge concerning mileage driven under different conditions is of assistance in predicting crash involvement. However, the cost of accurately measuring this variable is very high. In the study referenced, an attempt was made to control this effect by randomly selecting driver records. Thus, the effect of this variable is treated the same as unexplained error. It is important to identify as many of these error components as possible in order to make a judgement concerning biases that might be introduced by non-random treatment of these variables.

LOCAL CONTROL

In situations where the investigator strongly suspects specific intervening variables to bias project effectiveness, he may compensate by using local control in assigning the levels of the project to the experimental units. This procedure, which is commonly called "blocking," involves the random assignment of the treatment (project) levels to each level of the intervening variables. For instance, if we suspect that the variable "IQ" affects the use of simulators or the driver education effectiveness criteria, we could break our sampled populations into sub-groups based on levels of this variable, and randomly assign the simulator to 50% of each of these sub-groups. In this way we control the bias and variability which would be introduced if I.Q. were not controlled.

MEASUREMENT OF INTERACTIONS

It is important to recognize that there are many interactive relationships among variables which relate to subsystem effectiveness. Therefore the effect of a combination is often very different from the combination of individual variable effects. This may be due to the interactive nature of the measured variables themselves or the fact that these measured variables are functions of more fundamental unmeasured variables.

It is possible to estimate interactions by selecting appropriate combinations of independent variables at fixed levels and averaging the response over these. The details of this approach are outlined in any of several references on experimental design. (See for instance 4, 5, 6).

As we have indicated, because of our inability to hold independent variables at fixed levels, the accurate estimation of such interactions is in general much more complex. Multivariate statistical tools are necessary to unravel these more complex relationships.

5. EXPERIMENTAL CONSIDERATIONS PARTICULAR TO HIGHWAY SAFETY PROJECT EVALUATION

We turn now to a more detailed examination of those experimental problems which frequently arise in highway safety project evaluation. An understanding of these problems and procedures will provide OHSP with increased ability to:

1. Understand deficiencies in research studies which were designed to measure the effectiveness of particular countermeasure programs.
2. Evaluate proposals for experimental evaluation which have been submitted to OHSP.
3. Guide agencies in the development of experimental evaluation procedures for their specific projects.

RESEARCH STRATEGY

The inability to achieve control on intervening variables by presetting them at specified levels causes severe problems in attempting to attribute changes in a response variable to changes in specific independent variables. In this chapter we examine research strategies designed to regain some of this control. Specifically we conceive three basic types of project evaluation study.

TYPES OF STUDY.

Before-After Study. In this type of study the investigator performs analysis on the same population before and after the change in an independent variable is initiated. For instance, we observe a population of problem drivers before and after a remedial training course. Since the basic population characteristics are the same in both populations, conclusions are not masked by these intervening variables.

However, those characteristics which vary over time may introduce enough bias and variability to mask or exaggerate the true effect. For instance, consider the introduction of a stoplight at a problem intersection. This is a direct component change operating on the pre-crash environment. As such, an appropriate measure of effectiveness might be number of crashes in a fixed time period before and after the countermeasure. Although the physical environment is maintained during the study, the operator and vehicle components may vary. If in fact the signal diverts certain "problem operators" to other intersections, the measured reduction in collision rate does not reflect the true project benefit.

Another problem is introduced if the change causes a short-term transient response and the evaluation takes place in this period. For instance, in the above example, if due to habitual practice operators are unaware of the light initially, the collision rate may go up due to panic stops, failure to yield, etc. This short-term increase does not reflect the true project benefit.

Parallel Study. The parallel study involves the selection of control populations and active populations subjected to some change. The dependent variables are them compared in these populations to determine project effect. Here one must attempt to control population differences in such a way that any observed differences in the dependent variable can be attributed to the independent variable of interest.

In general the improper selection of variables to be controlled limits the usefulness of such a study. Suppose we wish to evaluate a stoplight by comparing crashes at two intersections. We "match" intersections on traffic flow, environment, etc. However, if the operators using each intersection differ in their risk-taking behavior, the true effect of the signal may be "confounded."

A Posteriori Study. This type of study is aimed at inferring the causal structure of some event or change that has already taken place. For instance, one might try to determine if the decrease in 1967 Michigan fatalities can be attributed to specific causes other than chance.

In this situation the investigator is free to examine a wide range of causal hypotheses and to exercise control by selecting sub-populations and variables of interest. However, because of the historical nature of the data, the quality and quantity of the data may be deficient.

The choice of type of study involves a priori consideration as to the types of variation which may confound experimental results. In addition, administrative and cost considerations influence the decision. In certain instances specific strategies are impossible. For instance, a before-after study of teenage driver education cannot use number of crashes as a criterion, for those who haven't taken driver education may not drive.

ELICITING CAUSE-EFFECT RELATIONS. We wish to evaluate project effectiveness in terms of a causal-chain relating the project to system objective. This induces a severe methodological problem, for the overrepresentation of some variable in a population does not necessarily imply that it is causative. For instance, the overrepresentation of the variable "speed" in freeway crashes may be simply due to the fact that "speed" is common to freeway driving. The overrepresentation of "women" in certain crash situations may be due to the fact that women are more frequently exposed to the conditions generating these situations.

This problem may be better understood by using some simple concepts of probability theory. We use the following notation:

Let

$P(A)$ denote the probability of some event A. If A is the event "accident", $P(A)$ is the probability of an accident.

$P(A|B)$ denote the probability of A conditional on the occurrence of the event B. If A is the event "accident" and B is the event "speed greater than 70 mph", the $P(A|B)$ is the probability of an accident given that speed is greater than 70 mph.

Suppose we are interested in determining the probability of an accident given that speed is greater than 70 mph. Then in the accident population we might measure the "frequency" of speeds greater than 70 mph. That is we might measure $P(\text{speed greater than 70 mph}|\text{accident})$. A high value of this probability does not necessarily mean that speed is causative, for $P(\text{accident}|\text{speed 70 mph})$ may still not be large. This can be seen from the following relationship:

$$P(\text{Acc}|\text{Speed } >70)$$

$$= \frac{P(\text{Speed } >70|\text{Acc})\ P(\text{Acc})}{P(\text{Speed } >70|\text{Acc})\ P(\text{Acc}) + P(\text{Speed } >70|\text{No Acc})\ P(\text{No Acc})}$$

Thus, in order to examine speed as a causative factor, we must look at the frequencies of speed in both the accident and non accident populations. For example, suppose that 95% of all vehicles in a certain accident class were driving faster than 70 m.p.h. Then $P(\text{Speed } >70|\text{Acc}) = .95$. However if $P(\text{Speed } >70|\text{Not Acc}) = .95$ and $P(\text{Acc}) = .05$

$$P(\text{Acc}|\text{Speed } >70) = \frac{(.95)\ (.05)}{(.95)\ (.05)\ +\ (.95)\ (.95)} = .05$$

The probability of an accident given that the driver is going greater than 70 mph is unchanged from the unconditional probability. This is true because the "occurrence" of speed is the same in both accident and non-accident populations.

Suppose that everything is the same, except that P(Speed >70|Not Acc) is smaller - say (.50). Now P(A|Speed >70) is increased to .0909. Again, notice that this is still much less than P(Speed >70|A) = .95.

Erroneous use of the latter measure results in conclusions that may be completely unrealistic. Measurement in a control population which does not display the characteristic of interest is essential in eliciting "causative" information on this characteristic. Summarizing, this section points out the dangers in inferring cause from the wrong data, or failure to select controls (non crashes), and of attempting to attribute accidents to simple, singular causative factors.

THE PROBLEM OF TIME. As we have seen, variation of some process over time may confound the variable of interest, especially in a "before-after" study. In addition, the parameter "time" may serve to hinder research in other manners. First, traffic crashes are rare events and, consequently, any inferential procedure developed to study them must be designed to obtain data over a substantial period in order to accumulate a sufficient sample. Unfortunately, other variables change simultaneously either in deterministic fashion (age) or in a probabilistic fashion (general traffic patterns). In addition, personnel and administrative policies may vary during the data collection processes. These may introduce undesirable biases or variability. Furthermore, subjects under study may drop out of the population thereby decreasing the sample size.

EXPOSURE. The problem of exposure--"identifying those circumstances present in accident cases in larger measure or more frequent degree than in the uneventful population of risk situations"* is a key factor in enhancing the meaningfulness of accident-oriented research. For example, it seems obvious that "teenagers" are overrepresented in the crash population; however, is it possible that this is because they are exposed in larger numbers or to greater "dosages" of night driving, random driving, etc? The answer to such a question is clearly of relevance in selection of proper countermeasures to reduce this overinvolvement. Again, quoting from Jacobs (7):

> Failure to recognize and deal with this problem has resulted in an unfortunate research situation. Analytical results which possess no more than speculative value are being constantly generated. Despite the seeming simplicity of these research problems, we still do not know whether men are safer drivers than women, whether it is more dangerous to cross the street with the light or against it, whether girls are stronger swimmers than boys, or whether aspirin is more deadly than lye. We do not know whether excessive speed is a factor common to turnpike accidents or common to turnpike driving. In short, there is a major problem in separating those circumstances which are associated with the occurrence of risk situations.

This problem is inherently related to the choice of a control group, whereby exposure is measured and controlled.

---

*Jacobs (7), page 332.

But we do not know what factors combine to measure exposure. Some investigators have assumed it is related to miles driven per year, or perhaps miles driven under certain environmental conditions. However, it is easy to make specific criticisms of such measures. We do not even know what "units" exposure is measured in, although some recent research suggests that ordinal measures are more appropriate than interval (8). In addition exposure is a probabilistic, not deterministic, phenomena. Thus, even if we can predict its expected values for a driver or section of roadway, for accurate evaluation, we must consider variation about this expected value.

CONSIDERATION OF COMPLEX ALTERNATIVES. It is our opinion that, at the present level of knowledge, sophisticated methodology and analysis can only reduce, not eliminate the number of possible explanations for a crash related phenomenon. Careful abstraction of sub-systems to more manageable parts and consideration of intermediate levels of performance may ease this problem, but we are sure they will not eliminate it. Hence, the investigation must be charged with the responsibility to explain other factors which might combine to yield the observed results. Careful consideration of a number and variety of complex causal effects leading to the same data structure will not only enhance a specific proposal, but will do much to improve the body of knowledge on the phenomenon by aiding and encouraging other research studies to explore and reinforce or eliminate these alternative explanations.

GETTING THE DATA

We have seen that verification and implementation of analytical models and development of observational hypotheses require obtaining and analyzing data. At this time, we shall discuss the former problem in greater detail.

CHOOSING THE POPULATION. Ideally, the population we select to obtain data from should correspond exactly to the population of interest--target population (Figure 10). For instance, if we wish to estimate driver characteristics in the State of Michigan, we certainly would not wish to get our data only on Wayne County drivers. Alternatively, if we wish to measure collision rates on I-94 in 1968, we would not necessarily make this inference on the basis of historical rates. In the first case, the sampled population differs from the target geographically*, while in the second, this difference is in time. Clearly, such differences weaken inferences, due to disturbing variables that are not measured and due to variability in the variables of interest that may exist between the sampled and target populations.

However, in many studies, non-overlapping sampled and target populations may be chosen with full knowledge of

---

*And probably in demography, economic status, and in many other characteristics.

differences and deficiencies. This may be due to several reasons:

(1) Convenience and ease of obtaining and verifying information in a specific population over which the investigator has familiarity, "logistic" control, and liasion with officials who can provide aid.

(2) Prior knowledge that the sampled population, although deviant in geographic location and/or time, does not differ significantly in characteristics of interest.

(3) Consideration of the costs in establishing a wide-ranging data collection program.

Unfortunately, many populations are selected and sampled with little consideration of the problems mentioned above. The failure to account (either explicitly or implicitly) for these has resulted in many research studies which are internally sound but which have very limited "extendability" over time and/or place.

CONCEPTS OF SAMPLING. Once the sampled population is selected, the investigator must develop a procedure for obtaining data from the population. In many cases, he can sample exhaustively--all units can be measured. For instance, all 1968 Michigan fatal crashes may be analyzed*. Alternatively, if the population is large or if costs of sampling

---

*Notice again that using these as a representative sample of all 1968 U.S. fatals induces the problems raised earlier.

are high, the investigator may wish to select some subset of the sampled population, measure the characteristics of the subset, and from these, infer the characteristics of the larger sampled population. In this section, we shall briefly indicate alternative ways of selecting this subset or sample.

In simple random sampling, (10, Ch. 2), we choose a sample such that every sample of size n has an equal chance of being selected. In practice, the sample may be constructed by drawing n. random numbers between 1 and N(sampled population size) and selecting units corresponding to these. If each unit can be selected at most once, we are sampling without replacement. If, on the other hand, each unit can appear more than once, we are sampling with replacement. The choice of a procedure influences our sample "estimates" of the population parameters.

Random sampling is probably the most widely discussed procedure for generating data. This is because the procedure:

(a) offers a relatively simple way of getting data.

(b) is the basis of almost all statistical models used in the analysis of data.

(c) generates population estimates which are "unbiased" and simple in structure (10, Ch. 2).

In practice, however, there are many data sets which are acclaimed to be or are analyzed as random samples from some population when they are in fact <u>not</u>. Although the effects of using such procedures on data analysis are not very well known, it appears that dependencies and deficiencies induced by careless non-randomness may confuse those interpretations the investigator desires to make.

One alternative to random sampling is systematic sampling. Here, the first unit of the sample is selected and thereafter every rth unit is picked until a sample of size n is generated. For instance, we might investigate owner characteristics of every rth car through an intersection. The two major advantages of this procedure are the ease in getting a sample and the possible gain in precision due to "spreading" the sample over the population of interest.* The major disadvantages are the inability to estimate this precision with confidence and the inability to detect periodic phenomena present in the population. Furthermore, there is little work done on using mathematical models to analyze systematic data--time series analysis is one notable exception.

A second alternative is cluster sampling or some variant thereof. In cluster sampling, we choose as a sampling unit some group of elements (a cluster). The selection of clusters for a sample may be made systematically or randomly, and clusters may be of equal or unequal size. Reasons for cluster sampling include failure to have lists of the population elements individually but only in clusters, (households, persons in vehicle, etc.) and economic gains from the convenience of getting data on larger groups at once. In cluster sampling,

---

*In random sampling, there is some finite probability that all sampled units come from a small sector of the sampled population.

the units within a cluster may exhibit dependencies, and these must be considered in the analysis. In general, the data analysis is more complicated and again, mathematical statistical models are not available to the extent of random sampling.

VARIABILITY. We have used the term variability without carefully defining and describing its meaning. For the purposes of this discussion, we shall consider three types of variability: (a) Variability between the target and sampled population (b) between the observed sample and the sampled population, and (c) within the sampled population itself.

Before proceeding, however, we must define what is meant by variability and central tendency. In any collection of units where the characteristic(s) of interest is not uniform, one is interested in the value which seems to be most representative of the collection. This search for central tendency may be made by inspecting a frequency plot of the data or by using appropriate summary statistics (mean, median, mode). In addition, one is interested in the spread of the data about this central point. This spread or variability may be observed or estimated by appropriate statistics (maximum deviation, standard deviation, etc.)

"Non-overlapping" target and sampled populations may induce differences in both the central tendency and variability of the data (see figure 10). The extent of these differences is not observable however; they must be based on the investigator's subjective opinions--either quantified or qualified. Reiterating an earlier opinion--it is these differences which may most seriously bias and reduce "extendability" of a research study.

The second source of variability--the difference between the observed sample and the sampled population--is largely controlled by the sampling scheme. Again, a random sample seems to be the most general method of controlling such variability although other schemes, either standard ones or ad hoc procedures, may be more appropriate in specific analyses. Again, the choice is dependent upon the investigator's prior notions. Possibly pilot studies may be conducted to supplement these.

Finally, we have within sample variability. This is the only observed variability of the three types we have indicated. In most situations, we use this variability, appropriately modified, as an "estimate" of the variability of the sampled population. The proper summaries and modifications of this variability depend on the sampling scheme, and a priori considerations as to the sampled population.

PROBLEMS IN GETTING DATA. A large variety of problems arise in practical implementation of any of the above procedures. In this section, we shall summarize these briefly.

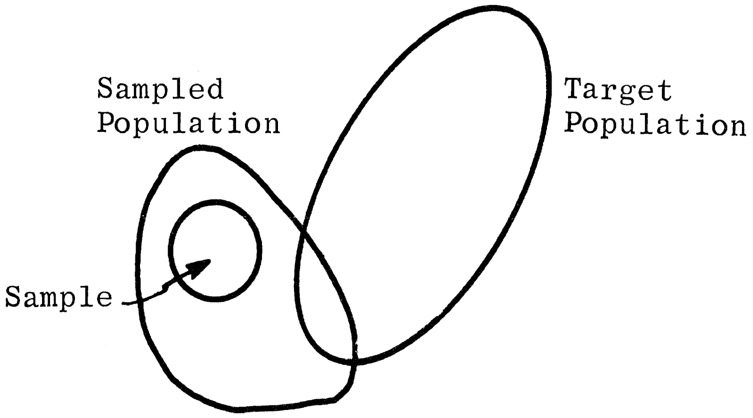


Figure 10. "NON-OVERLAPPING" TARGET AND SAMPLED POPULATIONS

Since each study contains unique problems, it is impossible to make general statements about the seriousness of each problem area and about appropriate procedures for their solution. However, failure to cope with these may easily result in erroneous conclusions and inaccurate data analysis.

Two related problems which commonly arise in obtaining information from persons are those of response bias and non-response. The failure to respond accurately and honestly or to respond at all (because of refusal, absence, etc.) may seriously hinder accurate conclusions. If these persons have characteristics which deviate from the remainder of the population, a bias in central tendency arises. Furthermore, the estimated sample variability may be either reduced or inflated, thereby altering the precision of the study.

Some positive steps are available to reduce the effects of this problem. Other data sources (neighbors, public agencies, etc.) may be probed to estimate the magnitude of the bias. This information may then be combined with the investigator's prior opinion to estimate and correct for the total effect. Non-response may be reduced by "callbacks" and by using other data sources. However, this is frequently expensive and inconvenient.

A second problem is induced by the investigator's "entry" into the system under study. People may drive differently if an observer is in the car, or if they know a research study is being conducted in certain locations. In countermeasure research, there is some evidence that merely asking persons questions during the study may alter their behavior thereby "confounding" any countermeasure effect. Cochran (10-11) has stated that any countermeasure that cannot change the system more than a questionnaire is probably of dubious worth. When the magnitude of the desired change is small and highly variable, the ability to distinguish these two effects decreases.

A third set of problems fall into the category of miscellaneous errors. For example, if responses require interviewer (data gatherer) judgment as to categorization, interpretation errors may easily result. A thorough training program and good communication between the investigator and the interviewers can minimize these errors. Additional bias and variability may be introduced through respondent error in recalling or describing certain variables. For instance, mileage driven in the last year, distance crash occurred from intersection, etc. are hard to estimate. Well designed, unambiguous questions can reduce respondent confusion and thus provide more concise results.

As we have indicated, general, complete approaches to these problems do not exist. Other than the references indicated in the text, it appears that ad hoc procedures have been developed to solve practical problems as they arise. Unfortunately, these procedures have not always been adequately considered in the highway safety literature.

THE PROBLEM OF CONTROL. We have seen that failure to measure and correct for certain types of variation (such as exposure) has resulted in serious deficiencies in many studies. The decision as to which variables to control in a particular study depends upon a priori consideration as to the extent and magnitude of their effect on the relationship of interest. Cochran (11) classifies such variables as (a) major variables for which adjustment is essential, (b) variables whose effects we would like to control but do not, (c) minor variables which are disregarded. Our discussion here hinges on methodology for initiating control over major variables.

Essentially, there are two ways to initiate this control of variability: obtain the data from various (pre-specified) strata or stratify the sample by matching on certain variables after the data is collected.

In stratified sampling, certain variables or categories are prespecified and the sample is obtained by selecting data from each of these categories. Two questions are of interest: a) what categories should be selected? and b) what sample size (within a category) is necessary? Answers to both require prior knowledge as to the within and between stratum variability as well as costs and ease in stratification.

In safety research an investigator might stratify his sample on age. He would then select categories of this variable (say 16-25, 26-34, etc.) and draw sub-samples of size n_i from each strata. If the variable of interest (number of accidents for instance) exhibits high between strata variation (as it seems to) then precision is increased through this procedure.

Post-sampling control may be exerted by categorizing a sample of size n into sub-cells which are determined by those characteristics of interest. [e.g., Age group categories might be selected and within each group the number of persons with 0, 1, 2,... crashes might be determined] The resultant categorization is called a contingency table. The investigator may subsequently be interested in comparisons of the individual cells--either in estimating the parameters or in estimating the between cell association. More likely, however, he may be interested in further examination of other characteristics as measured within a cell. Again, he has controlled undesirable variability by "matching" only those objects with similar, basic characteristics. The disadvantage of this procedure is that there is no prior control over the within cell sample sizes as in stratification. Consequently, certain cells of interest may be empty or contain insufficient sample sizes. Further, it seems obvious that the more variables we partition our sample into, the more likely this event will occur. Therefore, resulting sub-populations may become too small for meaningful analysis. A graphic example of this comes from Rapoport (12):

> Let us say that an insurance company wants to insure against the risk of a bank teller absconding with the cash. What odds should be given--that is, what premium should be charged for insuring against a certain amount? To answer the question one can take a sample of bank tellers that abscond with the cash, one can grade that sample in accordance with how much they abscond with, one can differentiate between married men and single men (these are reasonable differentiations; possibly married men would not abscond quite as easily as single men; or vice versa). One could differentiate between various ethnic background, various ages, and so forth. So, given a particular bank teller--let's say he is 30 years old, plays the flute, has 2 daughters ages 14 and 11, named Susan and Mary, is married to a woman 4 years his junior, etc; what is the probability he will abscond with the cash? If you put in enough of these variables, you can finally narrow the relevant universe down to a single individual. But in this "universe" there is no absconding rate, and hence nothing on which to base an estimate of a probability. This is true of any actuarial calculation. One is forced to choose between what to take into account and what to ignore, and these choices are made largely on a priori grounds. There is no assurance that they are the right ones, and this is why practically any statistical finding can be disputed.

However, as we have indicated, there are other variables--those that are known but not measured and those which are unknown--which influence the variable of interest. Additional research is necessary in order to determine these (or estimates of them) and to control their effects. At the present state, such research is in its infancy.

For example, suppose we wish to investigate characteristics of drivers who are similar "risk-takers". We might feel that age and mental ability are measures of "risk-taking" that must be controlled; hence, we "match" our sample on these measured characteristics. However, we may suspect that the social environment, basic behavior patterns, and other "intangible" variables are better measures of the variable "risk-taking". It then becomes necessary to quantify (or estimate) these and match our sample on the basis of these estimates.

SAMPLE SIZE. The decision of what portion n of the sampled population (size N) to select involves consideration of (a) the distribution of the variable of interest in the sampled population and (b) the incremental cost implicit in obtaining a unit of data. It seems intuitively obvious that the sample size n should be increased as the variability inherent in the sampled population increases, for we wish to obtain enough evidence to measure this variability and to control its effect on stability of our estimates. On the other hand, the optimal sample size should decrease with increasing cost per unit of data. Thus, the optimal sample involves a "tradeoff" between these two criteria.

In order to develop a more explicit decision rule for sample size, we must introduce quantitative prior information on both the population and on the cost function. The information on the population can be quantified in terms of "loss" function--a function that expresses the loss (in dollars, time, or some other variable) when an estimate based on

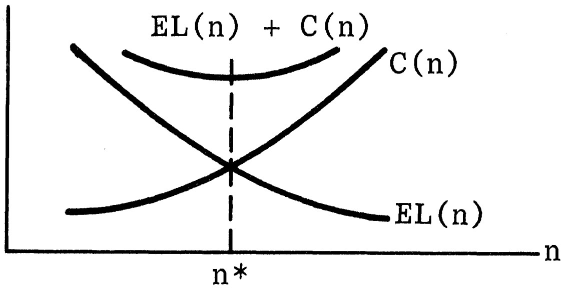


Figure 11. THE ECONOMIC SELECTION OF SAMPLE SIZE

the data denoted by $\hat{x}$ is used for some true value x. We denote such a function as L = f($\hat{x}$,x). We then average L over all possible experimental outcomes ($\hat{x}$,x), obtaining an "expected" loss EL(n) which is a function of the sample size n. If C(n) is the cost of obtaining a sample of size n, we then propose to choose n to minimize C(n) + EL(n) as in the Figure 11. Here n* is the "optimal" sample size. Again, specific selections of L(n) and C(n) can be made by utilizing more explicit prior knowledge.

DATA ANALYSIS

In this section we present a very brief introduction into the problems of analyzing data for meaningful conclusions. The complex multivariate situation relating independent variables to measures of highway safety subsystem effectiveness may require the utilization of advanced multivariate statistical tools in order to make accurate inferences. These tools will not be examined here. Instead we briefly examine some basic foundations.

UNIVARIATE ANALYSIS. Suppose we have a single measure of subsystem effectiveness denoted y. In addition suppose that our sample has been selected and properly matched so that changes in y can be attributed to some project. We now wish to measure the project effectiveness. To do this we sample the measure of effectiveness y in populations subjected to the project and in populations not subjected to the project*. Denote $y_1, y_2, \ldots, y_n$ the sample from the population subjected to the project and $y_1^*, y_2^*, \ldots, y_m^*$ the sample from the population not subjected to the project. If the project has increased subsystem effectiveness, the y's should be "larger" than the (y*)'s. Because of sampling variability the project may in fact be "effective" even if some y's are smaller than some y*'s. We are primarily interested in the project's effect on the "average" or "central-tendency" of the variable y.

MULTIVARIATE ANALYSIS. Under multivariate analysis a more complex and more typical situation will be considered. Specifically consider the situation where an effectiveness criterion has been defined and the change in effectiveness is related to the proposed project and several other influencing factors. For example the reduction in crashes at an intersection might be attributed to a new type of traffic light. In addition the reduction may be due to changes in weather, traffic flow, and the population operators. In this case the analysis cannot be made by simply comparing the number of crashes before the change and after the change. Instead the effects of other variables must be accounted for by using multivariate technique.

Multivariate analysis is a procedure which defines

---

*In a before-after study these populations are the same physical group after and before the project.

a number of states and obtains a measure of effectiveness for each state. Each state is described by levels of the independent or predictor variables. Consider the following states:

State	Traffic Signal	Traffic Flow	Variable
1	Old	Winter	factory workers going home
2	Old	Winter	housewives going to supermarket
3	Old	Summer	factory workers
4	Old	Summer	housewives
5	New	Winter	factory workers
6	New	Winter	housewives
7	New	Summer	factory workers
8	New	Summer	housewives

Many other possible state levels might be described depending upon the specific problem. These levels are developed by an analysis of the subsystem which the project affects. It may be extremely naive to evaluate the effect of a new traffic signal by comparing the subsystem response (accident rate) at states 1 and 6, for the effect of different traffic flows and different driver populations may in fact confound the measure of the signal's effectiveness. However, comparing states 1 and 5 provides an estimate of the signal's effect on factory workers in a winter traffic flow pattern. Many other examples can be easily constructed by the reader. A measure of total project effectiveness might be established by comparing states 1 through 4 with states 5 through 8.

Another important consideration in multivariate analysis is the form of the relationship between the variables and the measure of project effectiveness. The simplest form is an "additive" model in which the change in effectiveness is the same for changes in specific independent variables, regardless of the levels of the other independent variables. A more sophisticated relationship considers the interactions between two or more of the independent variables. For example a new traffic signal may reduce crashes more for housewifes than for factory workers. The key point here is the importance of carefully defining the levels of the independent variables. Consistent definition of the measure of effectiveness is of course mandatory for any analysis.

We anticipate that most project evaluations will fall into the area of multivariate analysis for the following reasons:

1. It is not possible to control the system completely enough to obtain a univariate measure of effectiveness.
2. The highway safety problem is a multivariate problem. It is desirable that projects designed to improve effectiveness be "robust" with respect to other independent variables existing in the system.

Many techniques have been developed for analyzing data from multivariate situations. The choice of a proper technique is directly related to the particular problem. It is important that the assumptions made by each technique are understood so that a correct interpretation of the results can be made. We recommend that the assistance of a qualified person be obtained if these assumptions are not readily understood.

It is important to keep in mind that these statistical techniques are tools which aid the analyst in uncovering and summarizing relationships. As such, results from their application should be tempered by a critical review of their assumptions and applicability.

For example, consider multiple regression analysis. If our response is a linear function of p independent variables, $y = a_1x_1 + \ldots + a_px_p$, this technique estimates the coefficients $a_1,\ldots,a_p$ from a sample of the data. Suppose that y is "reduction in crashes and x_1 is the level of vehicle inspection. In this case a statistically "significant" estimate of a_1 does not indicate that increased motor vehicle inspection reduces crashes. Multiple regression only establishes a linear relationship between two sets of variable. The extension of this empirical relationship to establish causal hypotheses is strictly the responsibility of the analyst.

6. CONSIDERATIONS IN COST-EFFECTIVENESS ANALYSIS

In the a priori evaluation plan of Chapter 1, measures of countermeasure and agency effectiveness are combined with cost information to develop a cost-effectiveness (cost-benefit) analysis of the project. The essential objective of cost-effectiveness analysis is to compare all potential project benefits with all project costs to determine the project value if it is implemented.

Cost effectiveness analysis has been successfully applied to the evaluation of large scale projects in the public sector. In this chapter, we shall survey some of the major techniques and principles of such analysis. The potentiality and limitations of applying these techniques to highway safety projects will then be discussed. Finally, we shall examine the specific problems involved in applying these techniques to each of the three levels of evaluation: evaluating individual projects, comparing projects in the same functional area, and comparing projects in different functional areas.

OVERVIEW OF COST EFFECTIVENESS ANALYSIS IN TRAFFIC SAFETY

THE COMPONENTS. Prest and Turvey, in their excellent survey of cost-benefit analysis (13), pose four basic components of the analysis: enumeration of cost and benefits, valuation of costs and benefits, discounting of costs and benefits over time, and consideration of relevant constraints.

Enumeration of Costs and Benefits.

(1) Benefits: We have argued that the ultimate benefits from a project are the improvement in system performance that can be attributed to the project. Unfortunately, this general evaluation criterion is of limited potentiality in practical project evaluation when the project involves an indirect component change.

In measuring the increase in phase objective attributable to a direct component change project we face the multivariate nature of each phase objective. Consider the pre-crash objective--reduction in potential for system breakdown. System breakdown is synonymous with crash. However, crashes are "rare events", and attempts to measure project effectiveness in terms of such events may induce severe sampling problems. Hence we might select "near misses" as our criteria of effectiveness, since the potential for system breakdown increases with these events. In this case severe definitional and measurement problems are also induced. Other measures might include changes in those variables that "correlate" with crashes--for instance, decrease in moving violations. It seems that the appropriate measure may be some mixture of these with other yet to be defined variables. The appropriate mix may in fact vary from project to project depending upon measurement and experimental problems.

Our crash phase objective is to reduce the severity of crashes. However, severity can be measured by several variables, including human injury or death, vehicle destruction, or environmental degradation. The appropriate mixture of these should be considered for an evaluation.

The post-crash objective has been stated as maximizing the appropriate treatment (medical and otherwise) at each defined post-crash time interval in minimum time. Here again "treatment" is a multivariate response and is highly dependent upon collision severity and the highly-variable condition of the post-crash system. On the other hand, attempting to optimize this phase by minimizing response and recovery time can yield sub-optimal results, for recent investigators have uncovered the importance of timely and effective medical treatment.

Two categories of projects within the "action" classification--informational and indirect component change projects--affect the phase objective in such a remote or intangible way that evaluation in terms of this ultimate objective is meaningless at the current state of knowledge. Hence the "causal" chain relating a projected change to system objective must be "cut" at some intermediate state for evaluation. The specific stage for evaluation depends upon knowledge of the functional component relations and the intervening variables in the chain. However, the evaluation should be conducted as "close" as possible to the phase objective in order to make subjective assessment of the ultimate change as accurate as possible.

It is of utmost importance to recognize that if the project does not affect the intermediate evaluation criteria, (provided these are correct), it will not affect the ultimate system objective. This results from the "causal" nature of intermediate criteria we have adopted. Hence failure to obtain an intermediate objective provides strong evidence for rejecting a project as a meaningful countermeasure. On the other hand, if the project does positively affect the intermediate criteria, it still may not affect the ultimate objective. Further analysis may be necessary to move down the causal chain from intermediate objectives to the ultimate objective.

It is our belief that various cost-benefit analyses conducted in highway safety fail to adequately consider the questions raised above. Recht (14), in applying C/B analysis to pre-crash countermeasures, chooses reduction in number of deaths, disabling injuries, and property damage crashes as his objective. However, the first two variables, deaths and injuries, are functionally dependent upon all three phases -- pre-crash, crash, and postcrash. Hence, changes in these are dependent on much more than specific pre-crash countermeasures. The Department of Health, Education, and Welfare (HEW) Motor Vehicle Injury Program (15) adopts the same inadequate criteria.

The second major deficiency in current C/B analyses, which the above studies also illustrate, is the failure to measure the effect of a project on a sub-system objective but instead to assume a certain effect. Measuring effectiveness is certainly the most difficult portion of an analysis. However, failure to do this forces reliance on "expert opinion". The latter phenomenon has certainly not resulted in an optimal allocation of funds. For instance, consider the functional area driver education. Recht assumes that a 1% reduction in crashes occurs for those who take the course. HEW (15) assumes a 0-20% reduction in injuries and deaths from 1968-72 based upon a massive effort to improve driver training. Yet, the HEW Secretary's Advisory Committee on Traffic Safety (16) has concluded "the present state of knowledge as to the effectiveness of driver education provides no certainty and much doubt, that the return on this enormous prospective effort* will be commensurate with the investment".

(2) Costs: The appropriate project costs to consider in a C/B analysis are the "opportunity costs" associated with the project. These are the costs of alternatives--other projects--that are foregone in allocating resources to a project. Most current C/B analyses associate the direct costs of a project with the opportunity cost. The direct costs are those resources expended in project implementation and maintenance. These may be further sub-divided into fixed costs and variable costs-e.g., those that vary with the intensity of project activity. In many cases these direct costs are appropriate; however, their blanket usage may lead to erroneous conclusions. Consider for example two projects yielding identical benefits but have different direct cost structures. Based on direct cost, if a selection must be made, we pick the cheapest project. Suppose however, that this project ties up the talent of a particularly able team of personnel for several years. Although the direct cost of the project (as measured in salary) is lower, the high opportunity cost of not using this team for other projects and programs may make this alternative less preferred.

---

*Estimated by the Committee at $142 Million in 1966 increasing to $330 million in 1972 if 100% national enrollment is achieved.

Valuation of Costs and Benefits. In general costs can be valued in monetary terms. Opportunity costs do not automatically have this property: however, if the alternative uses of the resource can be priced, these too can be valued in dollar terms. Therefore in this section we shall briefly consider the more difficult problem of valuing benefits.

Over the past few years a great deal of effort has been spent in attempting to measure the "cost" of highway collisions. (See for instance 17, 18). This can be undertaken by one of three methods. First, one might consider the direct costs necessary to return the system to normal capability. These include hospital costs, vehicle repair costs, and environmental repair costs. Second, one might add to these the opportunity costs of lost income during recovery, lost output to society during recovery, lost income due to death, and other imputed losses. Finally one might add another set of costs which impute personal and family losses due to death and injury, etc. Different criteria have been developed for measuring these latter two losses by Recht, HEW, and others. (14,15)

A similar problem arises in valuing intermediate benefits from informational or indirect component change projects. For instance, how does one value "increase in knowledge" or "increased information retrieval speed"? It is our opinion that these can best be valued in the units they are measured in. In the above examples, "change in test scores" and "decrease in clock time" may be appropriate criteria.

Traditional C/B analyses attempt to convert all benefits to monetary terms so that the cost benefit ratio is dimensionless. Then, a cost-benefit ratio greater than one suggests project rejection, while a ratio less that one suggests acceptance.

It is our opinion that such dimensionless quantities add little to the cost-benefit analysis of highway safety projects. This is true because the intermediate benefits which must be used to evaluate certain classes of projects are difficult to put in monetary terms and because the evaluation of ultimate benefits in monetary terms is at best subjective. The effect of these dimensional C/B ratios on project evaluation will be considered later.

Discounting Future Costs and Benefits. It seems reasonable that a project benefit "today" is worth more than the same benefit at some later date. Similarly a project cost today is "worse" than the same cost later in time. Because of this, future benefits and costs should be discounted to reflect their present value. Although this discounting is conceptually simple, the choice of a discount rate is a subject of much controversy in economics. Theoretically the discount rate should reflect the opportunity cost to society of reordering benefits and costs. However, measurement problems associated with this are enormous, for we do not know the components of such an opportunity cost and further, these are not uniform throughout society. Prest and Turvey present an adequate summary of this problem.

Consideration of Constraints. In conducting any C/B analysis, the decision maker faces a number of constraints which limit his ability to achieve very large benefits and/or very low costs. These include physical constraints bounding project effectiveness, legal constraints, administrative constraints, and budgetary constraints. In addition, uncertainty as to costs and/or benefits constrains the evaluation by defining a "safety factor" by which benefits must exceed costs to be acceptable.

The highway safety problem results from deficiencies in the ground transportation system. "Optimal" solutions to the highway safety problem may in fact be sub-optimal in the larger system. These broader objectives constrain project effectiveness. For instance, if we argue that speed is a cause of system breakdown, a blanket maximum could be placed on speed. However, such a maximum might increase congestion and hinder the flow of traffic to the extent that the overall transportation system is impaired. Consequently it is necessary to consider explicitly the overall objectives of the transportation system in evaluating project effectiveness. (Some of these larger objectives are listed in Hatry, (19) pp. 32-33).

We shall not discuss these constraints further; our point is that they should be recognized and considered when their effect is to limit project implementation and evaluation in some way.

APPLICATIONS TO PROJECT EVALUATION.

Internal Evaluation. We have emphasized throughout this discussion that the proper measures of project effectiveness may in fact be intermediate measures. For such a project we recommend that a "dimensional" cost-benefit analysis be undertaken for the reasons cited earlier. This analysis results in a set of estimated or inferred benefits and costs. In certain cases the cost-benefit ratio may be an appropriate summary of this set, although Crumlish (20) summarizes some arguments against this.

Since costs and benefits are often not in the same dimensions, no strict "reject-accept" decision rules can be derived. This is good, in our opinion, for it explicitly requires the funding agency to evaluate subjective, intangible aspects of the project. These include "state of the art" knowledge about effectiveness, evaluation of the project investigator, methodology, etc.

Comparisons within a Functional Area. It is quite likely that projects within a functional area have the same intermediate evaluation criteria. This may occur because of similarities in the "causal" chain, methodologies, etc. Therefore benefits will be in similar "units" for the projects, and direct comparisons may be attempted. However, intangible costs and benefits, as well as subjective feelings about ultimate project effect, still form a very important part of the analysis.

Comparisons between Functional Areas. This seems to be the ultimate goal of project evaluation, for we wish to allocate resouces to those functional areas with the highest payoffs. Unfortunately knowledge of the relative payoffs between such functional areas is in general non-existent. In fact, some researchers argue that these do not exist in the highway safety field. Instead a coordinated, multi-facted approach may be necessary. Quoting from (13), "the fundamental difficulty (in health projects)...is that of the multiplicity of variables--when there are manifold influences at work on life-expectancy, productivity and the like, how can one hope to sort out the unambiguous influence of a particular health programme or any other single causative factor?"

In making comparisons between functional areas on the basis of reasonable, measurable intermediate benefits, one must compare benefits having different measurement units. For instance in comparing driver education and motor vehicle inspection, one might be forced to compare, "increase in knowledge of driving principles" with "decrease in the number of defective vehicles." Here, again, subjective comparisons are the only way to resolve this problem at the present time.

SUMMARY. The conclusions in this section are mostly negative, for in the end, we believe that "subjective" project evaluation is needed. However subjective evaluation is not "random" evaluation; instead it utilizes the best assessments available of intangible costs and benefits. In addition, utilization of a careful cost-benefit conceptualization places such an analysis into a logical framework and forces the decision maker to consider a wider range of alternatives.

The fact is that traffic safety project evaluation must differ from the evaluation of many other health programs. In most health programs the causative "agent" is known or strongly suspected and often involves a single agent, e.g. polio virus. In the highway safety field, such causative agents are not adequately known or understood, and it is certain that they are multiple in nature. Furthermore, the severity of the loss is also a function of these agents and of the post crash phenomena. Thus, attempting to attribute reduction in system loss to individual projects is an extremely difficult task.

Consequently we must determine what intermediate benefits result from a project. If we can objectively detect these, subjective evaluation can lead to inferences on the ultimate project effectiveness. It is unfortunate that many of the subjective estimates made by researchers are biased with unverified optimism, for it seems to us that underestimation of ultimate project benefit is a less serious error than overestimation.

As research in this field progresses, we should be able to move from such subjective qualitative cost-effectiveness comparisons to objective, quantitative comparisons. This research will consist of filling in the "missing links" of the causal chain model we have used throughout this report. Hopefully, organizations such as OHSP can direct research in this direction by indicating areas where quantitative links are missing and by encouraging research in these areas.

INTRODUCTION TO APPLICATIONS

In the next four chapters we shall apply the techniques of Chapters 1-6 to the analysis and evaluation of four projects selected by OHSP.

These projects fall into various stages of the project chronology of Figure 1. Consequently, we shall use the projects to demonstrate principles specific to these stages as well as more general principles of evaluation.

The Livonia Streets and Traffic Project (Ch. 7) is one which is being considered by OHSP for funding. Thus, this project will be analyzed by considering agency problem definition, agency project formulation, and the a priori evaluation of these tasks by OHSP.

The Pontiac Driver Education Project (Ch. 8) will be analyzed by examining the problem definition and formulation. In addition, we will analyze the decision-making process concerning whether or not to conduct an experimental evaluation. Subsequent experimental design problems will also be considered.

The Michigan State Police ASMD Project (Ch. 9) primarily involves an analysis of the experimental evaluation necessary to measure the effectiveness of this countermeasure. A detailed experimental design is developed which will provide answers to the set of intermediate objectives under consideration.

The Michigan Department of State Driver Records Conversion Project (Ch. 10) gives an example of the ex post facto evaluation of a project. Hence, in a systematic way we compare planned objectives and costs with those which were actually attained. Additional benefits resulting from this project are also presented and analyzed.

7. LIVONIA STREETS AND TRAFFIC PROJECT

INTRODUCTION

In this chapter we consider the Livonia project. This project is an information seeking project--a special case of an informational project (Figure 7)--which proposes á study to define highway safety problems and to develop potential solutions to these problems.

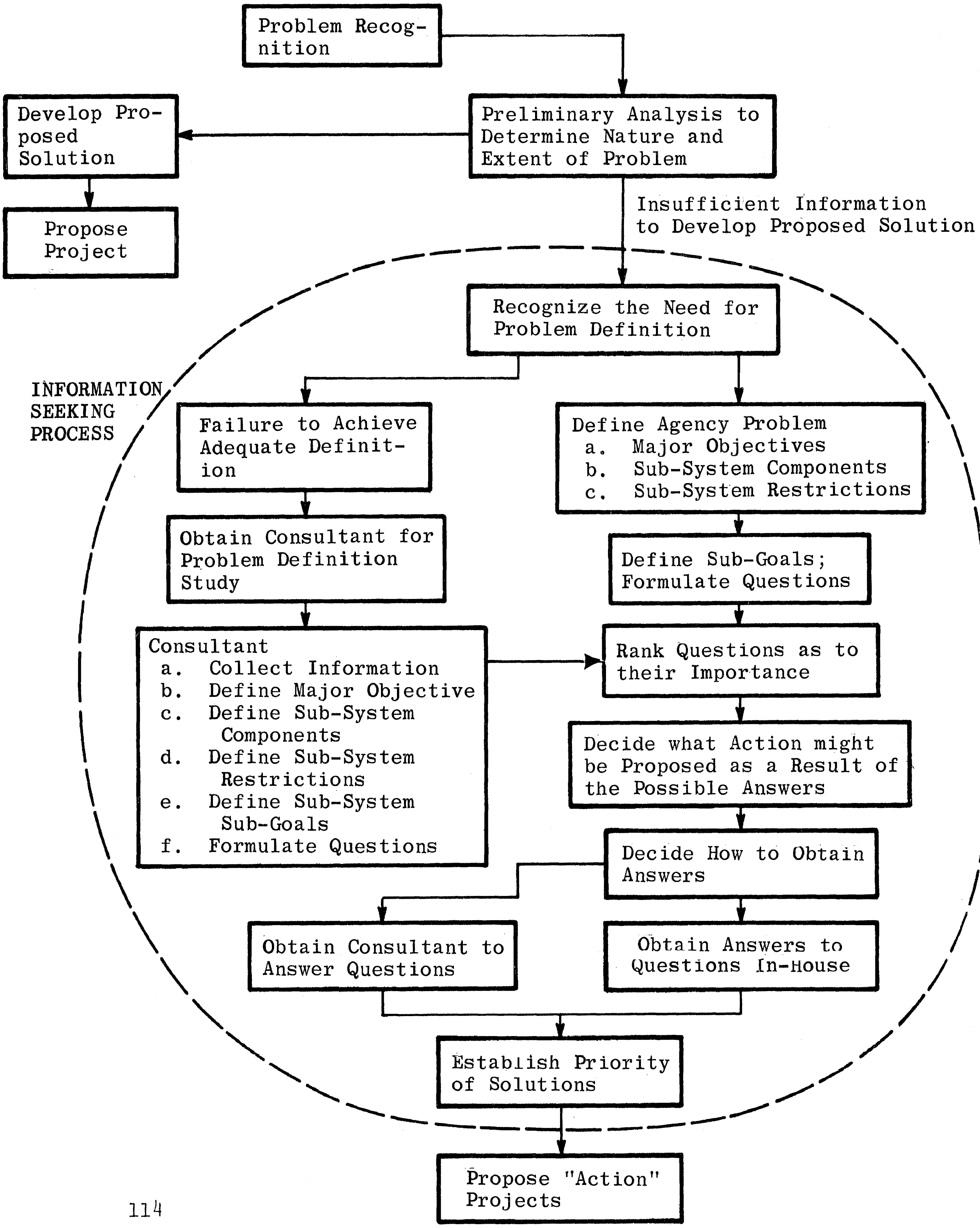
We feel that this project emphasizes the need for good project development (stage two of the project formulation chronology of Figure 1). In addition, the project provides information on how OHSP should conduct an a priori evaluation of information seeking projects. Consequently, our discussion of this project will be concerned with these two stages of the chronology. We will first present a procedure to be followed by agencies in developing a solution to their problem. This procedure will indicate the steps which lead to the decision to conduct an information seeking project. The OHSP a priori evaluation will then be discussed. This evaluation will consider in particular whether or not the agency has followed the recommended procedure.

PROJECT FORMULATION BY THE AGENCY

In this section, we shall develop an agency problem solving approach which will result in good project formulation. This approach can be described as shown in Figure 12. Initially, we have problem recognition followed by the per-

Figure 12

AGENCY PROBLEM SOLVING APPROACH



formance of a preliminary analysis. From this analysis may come either a proposed solution or the realization that further information is required before a solution can be developed. The Livonia project involves the latter case. In this situation, once the agency realizes the need for further information, it begins an information seeking process. This information seeking process could result in a proposal to OHSP for matching funds to help support an information-seeking project. In this section, we will indicate the steps that should be followed by the agency prior to deciding to make such a request.

The information seeking process should be guided by the principles of experimental design. This discussion will be developed from that point of view. Thus, it would be valuable to review the experimental design principles developed in Chapters 4 and 5. As stated previously, (Chapter 4):

> "The objective of experimental design is to obtain useful information with minimal effort. As such, experimental design is closely related to the scientific method--<u>proposing a hypothesis and the evaluation of this hypothesis by means of experimentation</u>".

Experimental design develops a logical structure which defines how measurements can be made to provide answers to a set of questions. Similarly, in this situation, an agency should develop a set of questions which can be answered by studies performed either by a consultant or in house. Thus, the studies performed are analogous to measurements made by an experimenter. Investigations which gather large amounts of data on the basis that it will be possible to analyze this by some method and extract useful conclusions are generally doomed to failure. By the same principle, studies which gather information must be well defined before the study is started.

The first step in the agency's "experimental design" or problem solving activity is a complete and concise statement of the major objectives. In short, what problems need to be solved? Statements such as: "Improvement of highway safety" are not definitive enough to provide guidance for constructing questions that can be answered by a study. However, an objective such as; "Reducing the number of intersection crashes", or "Improving the effectiveness of traffic signs" provide a basis upon which to develop a set of reasonable questions.

Once the major objectives have been established, it is necessary to define the components of the problem and the potential restrictions which may limit a solution.

The principles involved in doing this have been previously discussed in Chapter 3. In particular, the agency is:

> ". . .abstracting the essential sub-system components and parameters in a way so that project effects on the sub-system can be represented in a logical sequence."

This process should define what resources are available to the agency and what limitations are present in developing a solution.

For example, in the case of the traffic study proposed by Livonia, it is important to define the present road network, sources of traffic, and present traffic patterns. Potential restrictions on a solution include: (1) anticipated monetary resources available for implementing solutions, (2) extent of agency jurisdiction (e.g., the geographic boundaries of Livonia), (3) other changes being implemented that might influence the solution to the present problem (e.g., the effect of an expressway extension through Livonia on existing traffic patterns).

At this point, a causal chain which contains a number of sub-goals sequentially liked to the major objectives should be developed. For example, in the process of reducing the number of intersection crashes, the following might be established as sub-goals:

1. Inventory of the present intersections by usage, location, design, and collision rate.
2. Establishment of an effectiveness criterion for various intersections. This involves definition of the standards for determining the seriousness of intersection problems (e.g., number of personal injury crashes; total traffic flow.)
3. Study of the most critical intersections, using the principles of traffic engineering, followed by recommendations for intersection improvement.

A set of questions related to the alternative means for achieving these sub-goals can then be developed. The answers to these questions, will provide a solution or solutions to the larger problem. By this procedure, the agency can separate its problem into a number of small problems or questions.

The agency must then decide how it is going to answer each of these questions. These answers can either be obtained by members of the agency staff or by a consultant hired for this purpose. By considering the capabilities and time availability of its staff, the agency should quickly be able to determine which choice to make. Once this decision is made, the problem solver (consultant or in-house staff) can be given a well-structured problem in the form of a number of questions, along with the major agency objectives, the component definition, and the appropriate restrictions.

The extent and quality of problem definition by the agency will be strongly influenced by:

1. The information available to agency officials.
2. The capabilities of the agency officials.
3. The time agency officials have available for problem definition.

In certain cases, it may not be possible for the agency to perform an adequate "experimental design" or problem definition. In this case, a consultant should be engaged to perform a problem definition study (Figure 12). A slight modification of this would be regional study which defines objectives and problems for each of several local agencies. Because of the importance of good problem definition, OHSP should consider funding problem definition studies even though these are "one step" removed from projects proposing direct improvements to the traffic system. Such problem definition studies will provide a basis for specific highway safety recommendations.

The above procedure for agency problem solving is recommended by HSRI for the following reasons:

1. It provides a logical definition of the problem and its proposed solutions. Agencies using this approach operate under a discipline in defining their problem. For this reason, the agency can perform a better job of evaluating the information received from a consultant.

2. It provides for a better utilization of information seeking resources. Agency personnel should have an understanding of the agency structure, its goals, and its problems. Thus, they have immediate access to the information needed to structure their problem. Conversely, consultants generally have expertise suitable for solving particular problems. These are skills that would not generally be available to agency personnel. Through this problem structuring, the consultant's expertise can be applied to the problems it is best equipped to handle.

3. It equips the agency staff to implement the proposed solutions. A fundamental requirement in this procedure is the active involvement of agency officials and staff. As a result of this involvement, they will understand the development of the solutions. In addition, agency personnel will have participated in the development of the solutions. Thus, it can be expected that they will have a much greater incentive for implementing the resulting solutions.

EVALUATION BY OHSP

HSRI believes that the evaluation of information seeking projects by OHSP should consist first of determining how well the agency has adhered to the planning procedure recommended in the previous section. This evaluation can be divided into two major categories:

1. A priori evaluation of the proposal to determine whether or not the project should be funded.
2. Ex post facto evaluation to determine whether the project has been successful and to obtain knowledge that will be useful for evaluation of future projects.

The principles of evaluation are the same in both cases except that in the case of prior evaluation, OHSP will be attempting to predict whether particular goals can be met. On the other hand, in the ex post facto evaluation an actual determination of success in meeting goals can be made.

A PRIORI EVALUATION

In Chapter 2 we showed that informational projects must be evaluated on the basis of the quality and quantity of information produced. HSRI believes that the chances of obtaining useful information are greatly increased if the information seeking project is well designed. A well designed project will result if the agency applies the problem solving procedure developed earlier in this chapter. Specifically, OHSP must ask the following questions*:

1. Has the project objective been stated clearly and concisely?

---

*Another general question that might be asked in the case of engineering or planning studies is:

> What is the likelihood that the recommendations recommended by the study will be implemented?

If OHSP were certain that there was not any intention of implementing the recommendations, it would seem reasonable not to fund the project. However, it is unlikely that the answer to this question would be available. Thus, some evaluation of the proposer's motives is necessary. In addition, it would be reasonable for OHSP to develop a general policy in this regard.

2. Are the components and restrictions well described?
3. Does the proposal contain logically related subgoals and a set of questions to be answered?
4. What is the likelihood that useful answers can be obtained for the questions asked?
5. Is the information sought going to be useful in the solution of the agency's problem?

Questions 1 through 3 deal with the structure of the proposal and indicate how well the agency has assessed their problem.

Question 4 deals with an evaluation of whether or not the agency can obtain the information that it believes is needed to solve the problem.

Question 5 deals with the expected long-term benefits of obtaining the information sought by the agency. Thus, the answer represents an estimate of the ultimate success of the project.

In summary, HSRI recommends that OHSP's evaluation of information seeking projects concentrate on three criteria:

1. Has the requesting agency carefully analyzed its problem?
2. Can the desired information be obtained by the proposed project?
3. Will the information sought prove to be useful?

As indicated in Chapter 2, a scoring scheme could be established to aid in evaluating the trade-offs between these concentrations. For example, OHSP might decide that each of these criteria should be given equal value. In this case, a maximum of 10 points could be assigned to each criterion. If a particular criterion is satisfied completely by the agency's proposal, a score of 10 would be assigned. Lesser scores would be assigned depending upon the amount of deviation from the criteria. By this procedure, a "total score" could be assigned to each proposed information seeking project. These scores could then be used to rank information projects.

Alternatively, a minimum cut-off score could be established by OHSP that would indicate the minimum quality of proposed projects that is acceptable.

As an alternative, unequal weights could also be given to the three criteria by a different assignment of maximum number of points. For example, criteria 1 could be given a maximum of 10 points, criteria 2 a maximum of 20 points, and criteria 3 a maximum of 10 points. In this example, more value would be given to the potential usefulness of the information. Scoring schemes of this type are merely tools to help guide the evaluation process. They are not objective enough to provide final conclusive solutions.

This recommended evaluation procedure assumes that agencies do in fact follow the recommended problem solving approach. Thus, it represents a desirable goal. HSRI believes that the use of a logical problem solving approach, such as the one recommended, will greatly improve the value of information-seeking projects. Thus, it is recommended that OHSP take steps to encourage agencies to use this approach. This can be accomplished by:

1. The preparation and distribution by OHSP of a set of requirements that must be met by agencies who are submitting proposals for information-seeking projects. These requirements would be based upon the agency problem solving approach presented in this chapter.
2. Working with agencies to assist them in defining their needs through the problem solving approach.
3. Encouraging the use of consultants for problem definition studies when it is clear that these are necessary.

EX POST FACTO EVALUATION

As we indicated previously, ex post facto evaluation of information seeking projects follows the same basic procedure as preliminary evaluation. The same three criteria should be used after the project is completed to determine the actual project success in relation to these criteria. In particular, criteria 2--Has the desired information been obtained?--and criteria 3--Is the information useful?--can be evaluated directly after the project is completed.

8. PONTIAC DRIVER EDUCATION PROJECT

This chapter will be devoted to a discussion of how the project evaluation principles HSRI has developed apply to an indirect component change project--in this case the Pontiac Driver Education project. In particular, the first three steps of the evaluation chronology--problem definition, project development and a priori evaluation--will be presented together with a discussion of some experimental design principles applicable to certain sub-projects within the total project.

PROBLEM DEFINITION

The first step in the chronology is problem recognition and definition. As we indicated in Chapter 1, this step defines the direction of the remainder of the project. If the problem is not well defined, it is likely that the wrong problem will be solved. Therefore, in the a priori evaluation OHSP should consider whether or not the agency has adequately defined its problem. The Pontiac proposal (Item 6. 1b & c) indicates a need for an improved driver education program. In particular, the following deficiencies, noted in a 1965 study of Oakland County's traffic needs, are presented:

1. Driver education should be required one full semester.
2. Improved driver education practices are needed (such as the use of driver simulators).
3. Credit for high school driver education courses should be given.

4. More adult programs should be offered.

5. Driver safety schools should be established.

6. Qualified operators of school buses should be trained and periodically tested.

It is important for OHSP to note that:

1. The problem definition has resulted from a study conducted under the guidance of reputable organizations.

2. Specific problems have been noted as opposed to generalities.

From this problem definition a set of overall objectives or goals have been devised. These are listed below:

a. To insure that every eligible high school student and adult has an opportunity to enroll in a course of instruction designed to train him to drive skillfully and as safely as possible under all traffic and roadway conditions.

b. To provide a specialized program for Special Education students, physically handicapped students and adults, educationally deprived students, and violators.

c. To design a program that will not only develop the skills and knowledge required, but also to develop in each participant a positive attitude toward his obligations to all other users of public highways.

PLANNING

The next phase of the chronology is the development of a solution. In Chapter 1, we have stressed the importance of considering alternative solutions to the stated problem. The Pontiac proposal contains a well-detailed plan which includes specific areas for emphasis in Pontiac's proposed pilot program. The plan includes a time schedule with specific goals and it includes the assignment of responsibilities.

The major shortcoming of the proposal is the lack of well defined alternative solutions. Reasonable alternatives may not have been readily apparent and therefore were not included. Implicit in the proposal are two possible alternatives:

(1) Maintain the present driver education program.

(2) Make some minor improvements using locally generated funds.

Evidently the Pontiac personnel feel that this proposed project will do a better job in solving the problem than will the implicit alternatives mentioned above. Other alternatives to the Pontiac project will have to come from other agencies, and the comparison of alternatives will have to be performed by OHSP at the state level. Since this project is a pilot project designed to upgrade the driver education course, it might be compared with projects from other agencies which have similar objectives.

From the project plan it should be possible to extract a set of guidelines for conducting the project. For example, the Pontiac proposal contains a time table which indicates when particular phases are to be completed. In addition, specific sub-objectives are defined. For example: (from the Pontiac proposal, paragraph 6.3)

> "Development of positive attitudes in regard to the responsibility of driving a motor vehicle."

Upon completion of the project a comparison can be made between the guidelines for project conduct set forth in the plan and the actual project operation.

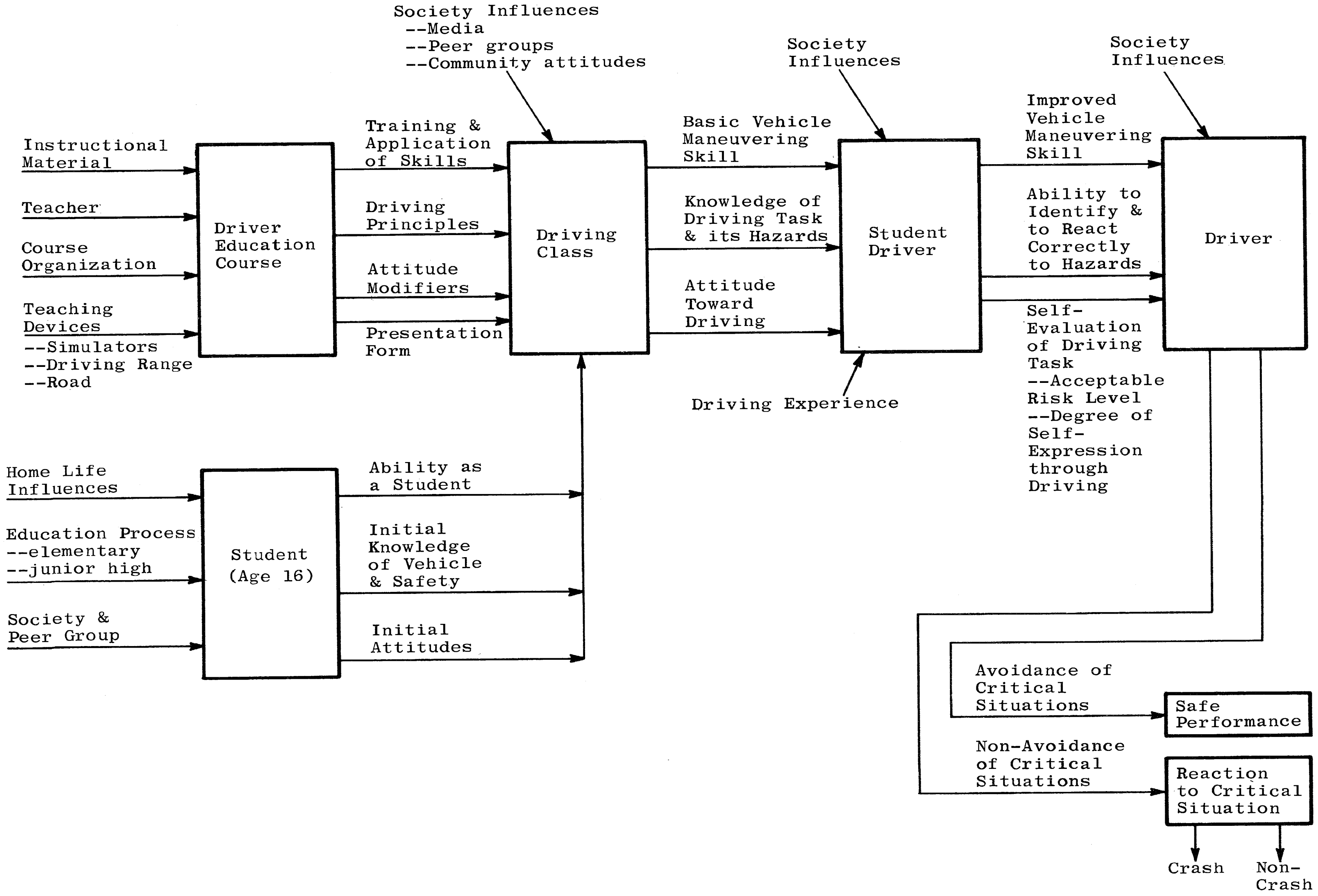
A PRIORI EVALUATION

At this stage a decision concerning whether or not to implement this project must be made by OHSP. The first step in this process is the determination---by studying the proposal and/or by personal contact---of whether or not the agency has effectively performed steps 1 and 2 of the chronology presented in Chapter 1. Namely, has the agency identified the problem and developed a plan for its solution. If this has not been done further information should be obtained from the agency. As we have indicated, study of the Pontiac proposal shows that these tasks have been accomplished. Applying the classification plan developed in Chapter 2, we can classify this project as an indirect component change project dealing with the operator, mainly in the pre-crash phase. Since it is an indirect component change project, we know that it will operate on the ultimate objective--reduction of number and severity of crashes--through a long causal chain.

A subsystem model showing the relationship between the various changes and ultimate objective is shown as Figure 13. From such a model the potential effect of this project can be evaluated. Decisions concerning the potential value of the sub-projects should be considered by studying their relationships in the causal chain. If it is decided that the

Figure 13

THE EFFECT OF DRIVER EDUCATION
ON THE DRIVER DEVELOPMENT PROCESS



proposed sub-projects will accomplish the stated objectives, OHSP should decide whether or not these stated objectives are worth the proposed cost. This decision will be influenced by the overall objectives that OHSP considers most important. In the case of the Pontiac project, the sub-projects appear to be related to the overall objective in a logical manner. For example, the elementary and junior high safety programs are designed to improve the students coming into the driver education class, and the introduction of simulators is designed to improve the quality of the driver education course.

A final evaluation question to be answered is whether or not the agency can do what it proposes to do. This question must consider the resources that the agency has available. In particular, the capability of the agency personnel is very important.

Based upon the criteria mentioned, the approval of the Pontiac project appears to have been a good decision. The one question remaining concerns possible alternatives. Alternatives were not mentioned within the Pontiac project, and it is not known whether alternative projects from other agencies have been considered. The selection of the best projects requires that alternative projects for achieving the same objective be considered. It is recommended that OHSP make such a comparison of alternatives a part of their evaluation procedure.

CONSIDERATIONS IN EXPERIMENTAL EVALUATION

We view our primary function in this section as providing assistance in answering the following two questions which are raised in Figure 1:

a) Does all (or part) of the project warrant experimental study and evaluation?

b) If so, how can we design an experiment to implement this evaluation?

Specifically, we have been involved in considering these questions with regard to three sub-projects: simulators, elementary/junior high safety, and remedial reading. These will be considered in greater detail below.

For each sub-project it is essential to examine the potential changes on the highway safety system in some logical way. Consequently, we have presented a simplified "model" of the driver education process as Figure 13. It is desirable to trace the effects of each sub-project through this model. In this way, reasonable intermediate measures of effectiveness may be chosen. In spite of the highly uncontrolled nature and influence of extraneous factors, we feel that the change in test score before and after the sub-project represents a reasonable measure of changes in knowledge, skill, and attitude.

We shall now expand on these concepts with regard to three specific sub-projects.

1. Simulator/Range Evaluation

a) Effect of simulators on attitudes.

It is fairly well known and documented that simulators can supplement traditional means for developing driving skills and knowledge. On the other hand, there is some controversy as to the effect of such devices on attitudes. It has been proposed that this effect be studied through controlled experimentation. The decision to do this must be based on the following:

a. What information on attitude changes and the effect of simulation on these is currently in doubt? Has other research answered this question?

b. Can an experiment be designed to measure this change (if it exists)?

c. Is this experiment worth the cost?

The general feeling throughout the meetings we have attended is that other research has not absolutely demonstrated positive effects of simulators on attitude change. Further the unique aspects of the three and four phase simulator/range programs have not been compared in any experimental way. Thus, an experimental evaluation of these programs seems worthwhile.

Our conceptualization of the general experimental process is shown in Figure 4. The explicit question in this case is whether the four phase program is superior to the three phase program.

The proposed experimental design consists of a parallel study (see Chapter 5) in which the four phase driver education course, which uses driving simulators will be conducted at Pontiac Northern High School. Concurrently, the three phase course will be conducted at Pontiac Central. This three phase course will replace simulator training with additional driving range training.

We wish to ascertain whether the change in attitude score is significantly greater for those subjected to the four phase program. In order to do this we must develop controls for other factors which could affect this measure of attitude. These factors include teacher differences and student differences. Teacher differences will be controlled by comparing only those students subjected to the same teacher. Student differences are uncontrolled in the present design. Consequently, biases may result because of different socio-economic factors at the two schools.

We recommend that an adjustment be made to compensate for any student variability that exists between the two schools. This may be accomplished in either of two ways. First, we might make an adjustment for each school by "subtracting out" the average difference in test score for each school as measured in the pre-test period. For example, suppose the average change in Central High for the pre-test period is 5, while the average change in Northern High is 8:

Suppose in the simulator evaluation we achieve an average change at Central of 5 and at Northern (where simulators are used) of 11. Then the change of 11-5 can be corrected for the initial differences by first subtracting 3 (8-5) from the Northern score.

A more sophisticated and accurate control can be obtained by adjusting each student's score by his grade point, IQ, or some other measure. This technique, sometimes called the analysis of covariance*, will remove student differences which are measured by these extraneous variables.

The specific test instruments being compared in the preevaluation period are the Siebrecht Attitude Scale and the Guilford & Schuster Driver Attitude Survey. Dr. Robertson's recommendation of the Guilford test, based upon its internal corrections and the quality of Dr. Guilford's work in human measurement, seems reasonable. The test, however, is constructed to be administered to a population of persons who currently are driving. Since it will be used on non-drivers in this case, we feel that revisions in either the test or in the test administration will be necessary.

A sample size of 192 persons is to be taken from each school. We feel that this is large enough for the problem at hand. Comparisons of change in test score can then be

---

*The analysis of covariance is a linear statistical model like those discussed in Chapter 4.

made for persons of the same sex, race, socio-economic level, and high school program to identify and control the influence of such factors on the study of student variation.

b) Use of simulators to measure Knowledge/Skill

Measurement of student driving skill is currently being made by a road test. In general, the Pontiac personnel feel that such tests are biased by instructor variation and student-instructor interactions. In addition, the traffic problems presented by a road test are not consistent. The simulator offers a potential source for removing such biases.

The possibility of using a test film with the simulators to measure student performance was considered in the meeting with Pontiac personnel. A study of Figure 13 indicates potential questions that should be answered in the process making a decision to replace the road test with a simulator test. The driver education course presents the student with principles that he integrates into his driving as he becomes more proficient. In using a test film and the simulators we are measuring whether or not the principles of driving have been transferred to the student. However, through a road test we are measuring the principles and the short term ability of the student to integrate these principles into his driving. Thus, the immediate objective of the course must be clearly defined.

Since the Pontiac officials generally agree that driving principles are being presented as opposed to creating qualified drivers, we recommend that simulators be used to supplement the current road test by utilizing a test film and by grading students on their performance. This will require careful consideration of the principles to be tested and the driving errors recorded by the simulators. The possibility of obtaining such a film from the manufacturer of the simulator should be investigated. Another possibility might be to splice together driving problems extracted from several different films. Comparisons should then be made to determine the nature of differences between simulator test scores and road test scores. If these differences are small (in both size and direction) the simulator may be able to replace the road test. On the other hand, examination of larger differences may indicate ways of improving the simulator program.

2. Evaluation of Elementary/Jr. High Program

This program will be initiated in all grades beginning in the fall. Since such a program is unique, there are valid arguments for conducting an experimental evaluation to measure its effect as students mature. Unfortunately, such an experiment is difficult (if not impossible) to conduct in any controlled fashion. This is true for the following reasons:

(1) The changes in skill, knowledge, and attitudes that occur as a result of the program are confounded (mixed up) with other changes which occur over time.

(2) Instruments which are useful in measuring such changes are constantly being revised, thereby making comparisons between time periods weaker.

(3) The effects of teacher variability and change in socio-economic level will be hard to control.

(4) The program is in its developmental stages. Hence, it is anticipated that many changes will be made as the program progresses. These changes could confound measures of effectiveness.

Therefore, we recommend that no experimental study be conducted on this sub-project. Instead, we feel that information should be maintained on the numbers and initial responses of those subjected to the program, as well as on the general direction of change.

3. Evaluation of Remedial Reading Program

An initial question which must be asked with regard to this program is: What are the objectives of combining remedial reading and driver education? These should be explicitly spelled out before any evaluation can be conducted. In our opinion, this has not been done.

It seems that the general objective is to improve the young driver's ability to read and comprehend driver education material. If this is true, then the program may be evaluated by administering a test designed to measure such a change before and after the course. Alternatively, a group of controls (persons needing the course but not assigned to it) could be used. Comparison of the test scores of the group receiving remedial reading with the control group (those not receiving remedial reading) provides a measure of the effectiveness of the remedial reading program.

On the other hand, possible program objectives include:

a) Change ability to read and understand road signs and signals.

b) Change driving attitudes of problem readers.

c) Use of driver education material to stimulate reading improvement in problem readers.

Since different evaluation instruments and experimental techniques may be needed for each of these, experimental evaluation is impossible without further consideration of the objectives.

9. MICHIGAN STATE POLICE ASMD PROJECT

In this chapter we will apply the methodology developed in this study to the Analog Speed Measuring Device (ASMD) project submitted by the Michigan State Police (MSP). This discussion will emphasize the requirements necessary for an evaluation of the effect of a project while the project is being conducted. Thus an application of the principles of experimental design will be presented.

The decision as to whether or not to fund the project should be made using the procedure presented in Chapter 1. In particular, OHSP should first look at the proposal and determine if the agency (in this case the M.S.P.) has identified its problem, established an objective and formulated a plan for reaching this objective. Once OHSP is satisfied that the proposal structure meets these criteria, it is then necessary to consider whether or not the objective is worthwhile.

Since the objectives of projects received by OHSP are varied, the first step in a priori evaluation is the classification of the project into its proper subgroup, using the plan developed in Chapter 2. The ASMD project is designed to implement--through purchasing and training--the usage of an improved speed detection device by the Michigan State Police.

Thus, it is classified as:

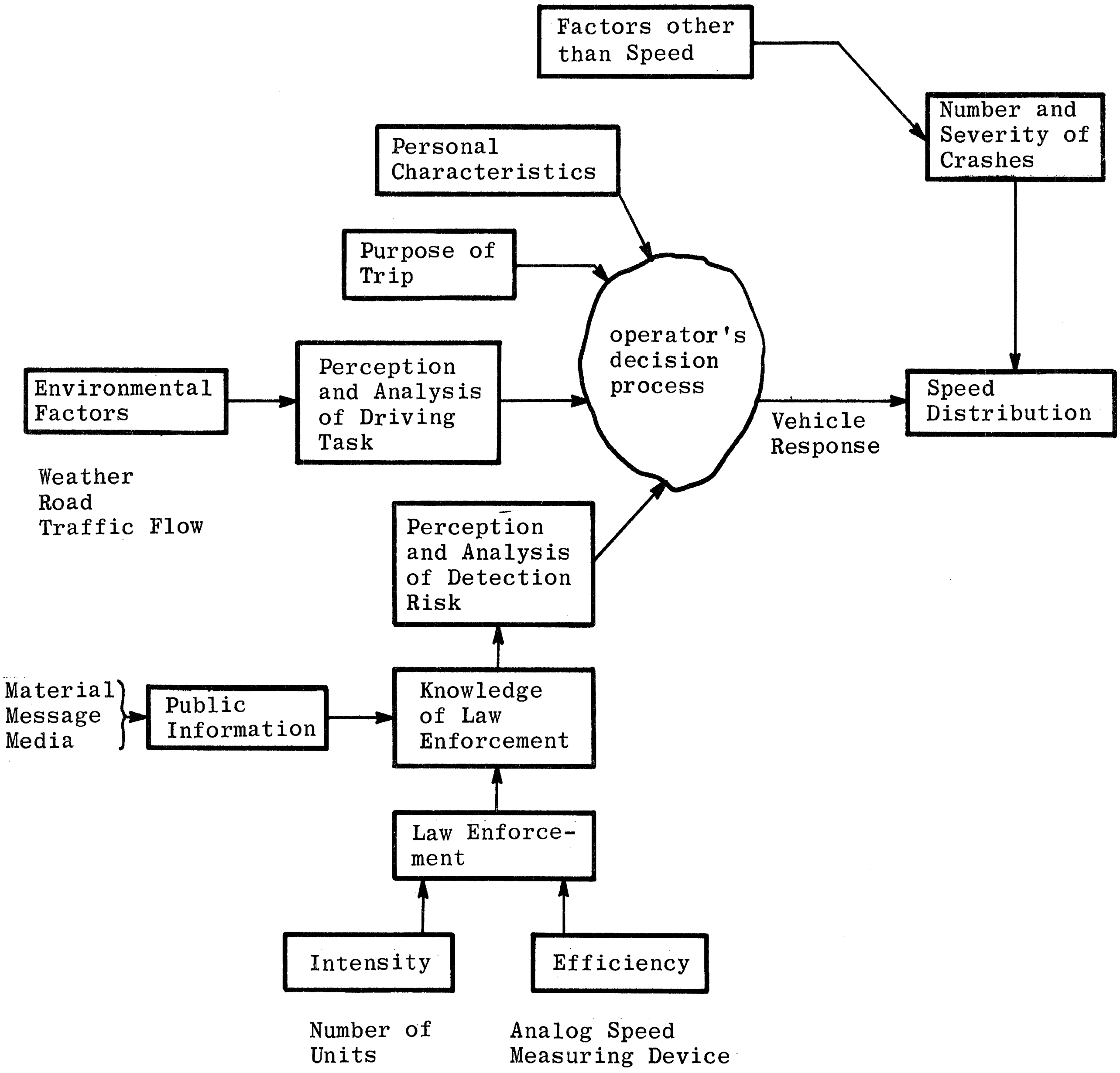
1. Indirect Component Change (from Chapter 2) "---effects are related to the system objective through a long, complicated and loosely defined chain of events---"
2. Both Pre-crash and Crash Phases
3. Operator Behavior Modification

At this point, a sub-system model should be developed in order to guide the analysis. Figure 14 presents such a model. The effect of this project--if one exists--is expected to operate on the causal chain by first improving law enforcement efficiency, followed by a modification of the operator's decision process. Hopefully this will reduce the number and severity of crashes. Each change in the causal chain must affect the following step if this ultimate objective is to be reached. The capability of this device for accurate speed measurement under various conditions when used by trained police officers has been established in several studies. (ref., 22 & 24). In addition, several studies have suggested that driver speed behavior is modified by more intensive (more units) law enforcement. (ref. 20 & 21). The laws of physics imply that higher impact speeds increase severity of crashes.

From this limited information an initial decision must be made concerning this project. We have evidence to indicate that the quality of law enforcement will be improved. We could ask the question: will improved quality of law enforcement modify the operators decision process so that he drives

Figure 14

SUBSYSTEM MODEL SHOWING THE
POTENTIAL EFFECT OF ASMD ON HIGHWAY SAFETY



at legal safe speeds? The limited information available to answer this question makes it unlikely that a definite answer can be obtained. Thus the initial evaluation of this project is limited to asking whether or not improved law enforcement is worth the cost of this project. If we believe that this improvement is worth the expense, it is reasonable to approve the project.

At this point, OHSP could decide to:

1. Approve the project
2. Perform further analysis
3. Reject the project

If OHSP decides to approve it, a decision might also be made to conduct a well designed experiment to evaluate the results of the project as it is implemented. Without careful evaluation of projects, progress toward the goal of highway safety improvement will be severely hampered. However, there are some important alternatives which must be considered. At one extreme, all available resources could be devoted to safety projects without any evaluation, while at the other extreme all resources could be devoted to a few projects, each of which includes an elaborate evaluation procedure.

We feel that some balance should be achieved between these two extremes. Each project should be analyzed using the principles presented in Chapter 1 of this report. This initial analysis or screening could be completed within several hours for some projects while others might take two or three days. A decision to accept, reject, or perform further analysis can be made as a result of this first analysis. Projects which are likely candidates for a detailed experimental study might also be identified by this initial analysis.

The following discussion presents a thorough analysis of this project, indicating an experimental design which is suitable for performing a detailed experimental evaluation of the effect of this project. The details of the design and the implementation of the study will require a significant effort. However, even if it is decided not to carry out the complete evaluation, this discussion indicates the thought process necessary to evaluate the potential value of this project and the problems that must be considered in attempting to perform even a casual evaluation of its effectiveness.

As we indicated previously, this project is an indirect component change which is designed to modify driver behavior. Potential improvements (if they exist) are expected to occur in the pre-crash and crash phase of the highway safety problem. Since this is an indirect component change, it is necessary to select some intermediate evaluation points at which to measure change. The points selected for this project are:

1. The level of public awareness of the existence and capability of ASMD.
2. The change in distribution of excessive driver speeds.

ANALYSIS OF PROJECT EFFECTIVENESS

In this section, the ASMD will be viewed in terms of its effect on one part of the highway safety problem -- the reduction in excessive speeding. To aid in the analysis of this effect, a schematic model of the subsystem to be modified has been developed, using the principles presented in chapter 3. Figure 14 presents this model. The key item in this system is the decision process of the driver, concerning, in this case, the speed at which he will travel. This decision process must be modified if excessive driving speed is to be reduced.

Although such a reduction seems to represent a valuable objective in terms of highway safety, we recognize that the ASMD introduction has many other potential benefits. First, it may allow the officer to continue with speed law enforcement in a more accurate, efficient, safe, and inexpensive way. Second, the introduction may improve officer morale by making his traffic enforcement role more challenging and interesting. Third, the use of ASMD may free some of the officer's time to conduct other tasks while on routine patrol. Fourth, improved accuracy in speed measurement may add increased consistency and efficiency to the judicial process. The evaluation at these peripheral benefits can be performed by accumulating and summarizing data from appropriate agencies throughout the regions under study.

These potential supplementary benefits can be classified under the heading of improved law enforcement. As stated previously these should be considered in the initial determination of whether or not to fund the ASMD project.

The remainder of this discussion considers the potential effect of ASMD intoduction of the driver's decision as to speed of travel. As shown in figure 14, some factors influencing these decision are listed below*:

1. Perception of driving task and response as influenced by:

    a. weather
    b. road geometry
    c. traffic flow
    d. type of area

2. Purpose of trip

    a. travel to work
    b. recreational travel
    c. professional driving

3. Personal characteristics of operator

    a. attitude toward law enforcement
    b. driving skill
    c. aggressiveness
    d. use of vehicle as means of personal expression

4. Perception and analysis of detection risk conditional upon a violation of the law.

The relationship between the introduction of ASMD and excessive speeding can be conceptualized in terms of a causal chain leading from the initial change to the final result. As is shown in Figure 14, the introduction of ASMD must first

---

*In addition see references 21 and 22.

have an effect on the operator's perception of the risk of detection when he speeds. Hopefully this effect will alter his driving behavior.

The intial effect of ASMD is expected to be an improvement in the efficiency of the existing level of law enforcement. Efficiency refers to the ability of each patrol unit to detect and document violations of the law. This is contrasted to the number of patrol units, which we have defined as the intensity of law enforcement*. The efficiency of ASMD has been established both by the Michigan State Police and by the University of North Carolina. (ref. 23, 24) This improved efficiency may introduce the following changes in the causal chain of Figure 14:

1. Improved efficiency of speed detection.
2. Improvement in overall law enforcement capability.
3. Increased public awareness of improved law enforcement.
4. Modification of the operator's assessment of his detection risk.
5. Modification of the operator's decision process.
6. Change in vehicle speeds which are in excess of the established limits.
7. Reduction in number and severity of crashes.

---

*Studies to measure the effects of increased intensity of law enforcement have been reported in (21) and (22). Although the experimental procedures followed in these have been carefully planned, the chosen measures of effectiveness may be highly biased by uncontrolled factors.

Thus there are a series of <u>intermediate</u> objectives which must be achieved before the ultimate objective -- reduction in number and severity of crashes -- is satisfied.

Since an important link in the causal chain is increased awareness of the improved law enforcement, HSRI feels that the effect of a controlled public information campaign should also be evaluated in connection with the study of ASMD effectiveness.

By studying the causal chain, a choice can be made concerning possible measures of effectiveness. This decision is influenced by a tradeoff between three criteria:

1. The measure of effectiveness should be closely related to a measure of the ultimate system objective -- reduction in the number and severity of crashes. Unfortunately, measures which are closely related to this objective generally are influenced by a large number of uncontrolled factors. In addition, because of the long causal chain, these measures may be insensitive to changes in the countermeasure under study. Hence intermediate measures where such contamination and insensitivity are avoided are desirable. The measure of effectiveness should be closely associated with the change being made in order to minimize the

effect of intervening variables.

2. However, an improvement at any step in the causal chain may not result in an improvement in the next step. Thus it is possible to achieve an intermediate objective without subsequently reaching the ultimate system objective.
3. The measure of effectiveness should be capable of accurate and economical measurement in the population of interest.

Based on the above considerations we recommend measurement at two points in the causal chain. These are:

1. The public awareness of improved law enforcement.
2. The distribution of vehicle speeds -- in particular the number which exceed the speed limit.

Thus we are interested in whether or not drivers are aware of ASMD and secondly, what action they take as a result of this awareness. If ASMD has an effect on excessive speeding, drivers must first be aware of its presence. If factors other than this awareness are controlled, then it is reasonable to attribute changes to speed to the improved law enforcement resulting from ASMD. On the other hand, a reduction in excessive speeding which is not preceded by increased public awareness represents a change due to other sources. Hence, measurement of change in public awareness serves as a valuable means for preventing erroneous inferences.

It is recommended that a series of "before-after" studies be conducted to measure changes in public awareness and excessive speeding. The following table summarizes the possible results of the measurements and the resulting conclusions (assuming other factors are adequately controlled):

Public Awareness	Distribution of Excessive Speed	Conclusion
Increased	Reduced	Drivers are aware of ASMD and they have modified their behavior.
Increased	No change	Drivers are aware of ASMD, but they have not modified their behavior.
No change	No change	Drivers are not aware of ASMD, and they have not changed their behavior.
No change	Reduced	Something other than ASMD has caused drivers to change their behavior.

In order to measure the single and combined effects of ASMD and of a public information campaign, it is necessary to select four regions in which the following changes will be made:

1.	Install ASMD	Conduct public information campaign.
2.	Install ASMD	Do not conduct public information campaign
3.	Do not install ASMD	Conduct public information campaign.
4.	Do not install ASMD	Do not conduct public information campaign.

Within each region before and after measurements will be made to determine the magnitude of any changes that may occur.

This four-region study will provide comparisons of speed and public awareness in ASMD and non-ASMD areas, both with and without a public information campaign. Thus, any observed changes can be associated with ASMD, with public information, or with the combined effect of these. We believe that the combined effect of both factors may be different from the sum of the individual effects of each factor. This design will allow us to measure such an interaction. In each of the regions measurements will be made prior to the introduction of any changes and after the changes have taken effect. Thus, the <u>change</u> in public awareness and excessive speeding will be measured in each region. Measuring the change will help to eliminate differences between locations which could bias the results.

Four locations which correspond to the State Planning and Development Regions (25) were selected in conjunction with Mr. John Longstreth of the Planning & Research Unit, M.S.P. The locations of these regions are shown in Figure 15. These regions are:

	Principal City	Treatment
I. Region 5 Shiawassee County Genesee County Lapeer County	Flint	ASMD & Public Information

Figure 15. STATE PLANNING REGIONS

		Principal City	Treatment
II.	Region 2 Jackson County Hillsdale County Lenawee County	Jackson	ASMD only
III.	Region 8 Oceana County Muskegon County Newago County	Muskegon	Public Information only
IV.	Region 4 Van Buren County Berrien County Cass County	Benton Harbor	No change

These regions were selected because of their natural separation with regard to newspaper circulation, commuting, and minimum traffic flow between regions. The objective in region selection was to minimize the influence of a treatment in one region on a treatment in another region.

It is recommended that in the two areas of ASMD treatment a maximum number of vehicles be equipped while in non-ASMD areas no vehicles be equipped. In this way, the maximum difference in ASMD utilization will be achieved, and maximum opportunity for detecting measurable differences is provided. If excessive speeding can be controlled by ASMD with the present intensity of law enforcement, measurable differences should appear by following this procedure. In this experimental design the effects of the treatments are compared in parallel. That is the treatments are started at the same time in each of the four geographic areas.

In order to obtain valid results from this evaluation if is important to control the other factors which influence driver speed and driver awareness. Thus, a major task in this evaluation is the establishment of a good experimental design procedure for conducting measurements in each of the regions. We will now indicate how HSRI proposes to design this critical portion of the evaluation.

DETAILS OF EXPERIMENTAL DESIGN. We have previously argued that the causal chain which relates the introduction of ASMD to changes in pre-crash and crash measures of effectiveness is highly affected by large numbers of intervening uncontrolled factors. Hence, intermediate effectiveness measures must be selected. The two measures we have recommended are: changes in public awareness of ASMD implementation within a region and subsequent changes in the distribution of vehicle speeds -- particularly those over the posted speed limit.

In this section we consider the experimental procedures necessary to obtain measurements and to compare these effectiveness measures in the four regions under study. We will draw heavily on the principles presented in chapters 4 and 5. Ideally, this experimental design must control those factors influencing the driver decision process which may mask or exaggerate the effects our experiment is attempting to measure. Since all of these factors are not known, and since controlling all undesired factors in the operational environment is impossible, any experimental plan must be tailored to control enough carefully chosen factors to insure the validity of conclusions we may draw.

ALTERNATIVE EXPERIMENTAL PROCEDURES

Three alternative experimental procedures may be considered to accomplish the goal mentioned above. Ideally, we might select a sample of operators in each area and measure the awareness and speeding behavior before and after the introduction of ASMD. This experimental plan controls the variability in characteristics within each region. In addition, by selecting operators in different regions with certain matching characteristics, comparisons between areas could be made. Unfortunately such intervention in the system is impossible without giving the operator knowledge of the experiment. Providing the operator with such knowledge will certainly alter his behavior and therby bias the results. Hence, this plan will not be considered further.

A second experimental procedure involves the selection of fixed road sites within each region and the measurement of speeds at these sites before and after ASMD is introduced. Independently, a sample of operators would be taken to determine the level of public awareness of ASMD within each region. Using this experimental plan we would select locations to control certain factors (e.g., road geometry, traffic flow, and purpose of trip) influencing speed. By selecting locations in each region with nearly equivalent characteristics, comparisons between regions would be strengthened. The problem with this procedure is that we do not have measures of the _change in_ awareness and subsequent _change in_ behavior for the same set of persons. Hence, we must infer that a change of awareness in one group followed by a favorable change in behavior in another group in the same region does in fact represent a regional change. The four proposed treatments may then be evaluated by comparing the magnitude of changes in awareness and changes in speed behavior in the four areas. As we have pointed out earlier, by comparing changes in these quantities, a portion of the regional bias is removed.

A third procedure also involves the selection of sites, using the same criteria as used in the second procedure, and the measurement of speed at these sites. In this plan, however, those vehicles whose speeds have been measured would be identified (by stopping them on the scene after speeds have been measured, by recording the license numbers and contacting their operators at a later time, or by some other procedure). In this way we might determine awareness and behavior for the same set of operators. However, since we have exercised no direct control in selecting the drivers to be measured, and since a different sample of operators will be selected in the before and after periods, it is possible that a third uncontrolled factor influences both awareness and speed behavior. For instance, the variable "educational attainment" may be related to both increased awareness of sub-system changes (ASMD) and to operation at legal speeds. In this case, it would be inaccurate to attribute this behavior to ASMD. Because of this possible bias, because on-the-scene operator identification is a difficult task, and because contacting specific operators is in general expensive, this plan also has some shortcomings. On the other hand, the ability to obtain measures of awareness and speeding behavior on the same set of persons is desirable in eliciting any causative relations which may exist.

Thus a final decision between procedures two and three will require a more extensive evaluation of their relative advantages and disadvantages. An initial effort in this study will involve the selection of one of these two experimental plans for detailed development.

THE CHOICE OF VARIABLES TO CONTROL

Both of the plans under consideration involve the selection of speed measuring sites within regions. We wish to select these sites both to control variablity introduced by the road and to insure that interregional sites yield comparable data. In general, we propose to select roads in each area by controlling physical characteristics, such as number of lanes, frequency of curves, and surface type. Traffic flow characteristics to be considered will include average traffic volume, density, and the average nature of travel over the road. Because of the importance of accurate site selection on the experimental conclusions an extensive effort will be made to determine exactly which site variables must be controlled. The selection of these sites will require map surveys and discussions with state highway department and state police personnel in order to determine specific locations which have the desired characteristics. Final selection will probably require decisions concerning the tradeoffs between ideal variable levels and those which can be measured with some precision.

In addition, it is necessary to control the disturbing influence of changes on the chosen road segments occuring over time. For instance, measurements should be taken only on those days and during those times when the controlled characteristics are close to the average values. It is also necessary to take data only when traffic is moving freely to remove the dependencies introduced by vehicular interaction.

The influence of weather will be controlled by making speed comparisons only for similar weather conditions. In order to maintain a reasonably small sample size, our tentative conclusions are that data should be taken only under good road and weather conditions.

Neither of the experimental plans we have presented offers a strong control on operator variability, although some control is maintained by selecting sites having similar average travel patterns and missions. We recognize that the regions selected for study have somewhat different socio-economic and ethnic characteristics. Within each region, it is also reasonable to assume that the characteristics of operators may vary from site to site. However, without the ability to feasibly and economically measure one set of operators throughout the period, it seems reasonable to control this variable only by obtaining large enough samples so that the average population characteristics are present, and to measure any change in effectiveness with respect to this average.

Additional control must be exercised over changes in public factors other than ASMD. For instance, enforcement levels and officer deployment should be kept constant at existing levels by all jurisdictions whenever possible. Other safety campaigns should not be initiated. Proposed route changes and construction should be avoided by selecting measurement sites where these events are not planned during the study period.

EXPERIMENT TIMING AND SAMPLE SIZE

In both experimental plans we are considering, before and after measurements are proposed in each of the four regions. Since we assume no public awareness of ASMD prior to its introduction, only speed distribution must be measured in each region before ASMD. We anticipate such measurements will take from one to two months depending upon the availability of State Highway Department crews and weather conditions.

After the introduction of ASMD the key question is when to begin measuring, for there certainly is a time lag between ASMD introduction and public awareness. (We hope that there is no lag between awareness and change in speeding behavior.) Figure 16 qualitatively depicts the problem.

It seems reasonable to make measurements after the transient, initial changes have occured. That is, we wish to make evaluations in the "steady state". Unfortunately, the time to reach this is not well known, and it may be different when a public information campaign is undertaken. A two or three month period may be adequate to allow these transient effects to dissipate. In addition, it is well known that awareness begins to drop at some later time if it is not reinforced. Hence, it would be interesting (although not essential to the evaluation) to conduct a second set of measurements (say after six months) to study this phenomenon.

Determination of the sample size necessary to ascertain the distribution of vehicle speeds and levels of public awareness within an area involves consideration of both population

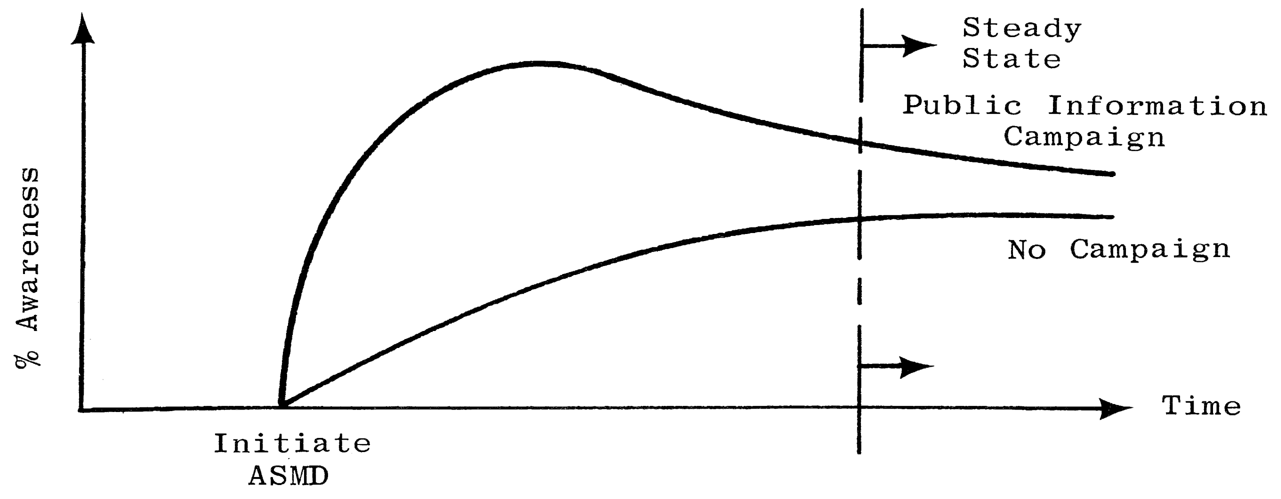


Figure 16. AWARENESS OF ASMD AS FUNCTION OF TIME

variability and cost of obtaining data. In our experimental plan several factors are controlled in making speed measurements; hence the appropriate sample size may be approximated by considering these.

For instance, if we measure speed distribution on high and low volume, residential and commercial sites within each region under good weather conditions for non-weekend traffic, (e.g., four locations per region), and if we require a 100 vehicle sample at each site, 1600 observations are needed per experimental replication. To control experimental error and errors due to other sources, two or three replications of the experiment are necessary. Assuming three replications, 4800 measurements are necessary in each of the before and after periods. Assuming that 50 vehicles/hour pass each measuring site, 192 hours of speed measurement time are required.

Since we are interested in knowing the proportion of persons aware of ASMD, our sample size for this sub-experiment should be designed to obtain an optimal estimate of this. The cost of obtaining information is highly dependent upon the interview process selected, so the necessary sample size will be selected after this process is chosen. In general proportions can be estimated with high precision using reasonably small random samples.

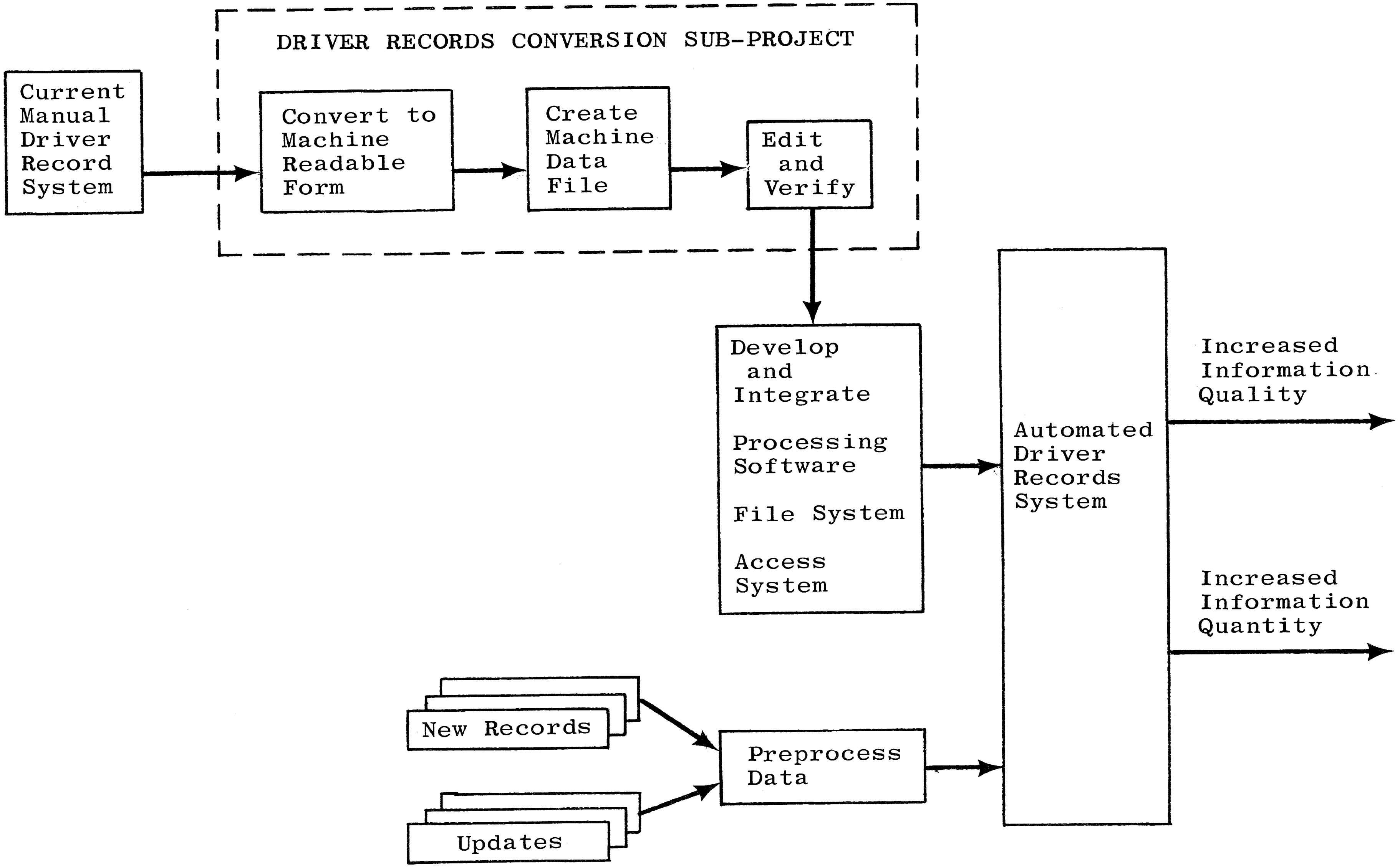
10. MICHIGAN SECRETARY OF STATE DRIVER RECORDS CONVERSION PROJECT

The Michigan Department of State (MDS) Driver Records Conversion Project provides an application of our ex post facto evaluation principles. Since this project is a sub-project in the development of a state-wide automated driver records system, we must consider the overall system in its evaluation.

Using the classification plan of Chapter 2 we observe that the automated driver records system project is an informational project involving the driver in all phases of the crash process. Consequently, this overall project should be evaluated in terms of the increased quality and/or quantity of information provided. However, in evaluating the conversion sub-project it is first necessary to develop a subsystem model so that reasonable intermediate measures of effectiveness can be selected for this sub-project. One such model is presented in Figure 17.

From this figure we see that the conversion sub-project is only one portion of the automated records system project. Hence, it is unreasonable to evaluate the sub-project in terms of the overall objectives (increased information quality and quantity), since many factors in addition to conversion influence these objectives. A more reasonable approach is to evaluate the sub-project by considering whether it contributes to the overall project objectives in a cost-effective way.

Figure 17.
AUTOMATED DRIVER RECORDS SYSTEM PROJECT



This is essentially the Department of State's Office reasoning in specifying an intermediate sub-project objective--to implement a portion of the record conversion process in an efficient manner.

In conducting a cost-effectiveness analysis of this type we are forced to compare alternative conversion procedures. Two such alternatives have been analyzed by the Department of State's Office--keypunching of records followed by machine file construction and typing followed by optical scanning to create files. A complete cost analysis of these by MDS showed that the second alternative is less expensive by $220,000. This shows the benefits that can be obtained by comparing alternative solution procedures in the project planning stage.

As we indicated in Chapter 1, an ex post facto evaluation by OHSP should be conducted to ascertain whether the agency achieved its objectives within the projected costs. In addition, any deficiencies which occurred should be analyzed.

The total estimated cost of the entire conversion subproject is 2.73 million dollars of which Federal funds paid $139,999 (5.1% of the total). It was estimated that 17 million documents for 4.5 million operators would be converted. The average estimated conversion cost is then $.16/document or $.61/operator.

Our conversations with personnel of the Department of State's office revealed the following information. The project is on schedule (53.8% of all files have been converted). This is true even though there was a delay due to the Federal lag in allocating funds. The total cost estimate of 2.73 million dollars is still current with seven months remaining in the conversion project.

In addition, MDS personnel have discovered that there are actually 5.5 million operators in their manual system. Consequently, the attained average conversion cost is likely to be $.50/operator.

Thus, we can say that the total project is meeting its stated objectives within the project costs. Therefore the Federal funds used in this sub-project were utilized effectively.

There are other unstated objectives which the project has attained. OHSP should also examine these in an ex post facto evaluation. The number of personnel involved in record look-up has been reduced from 98 to 55. The employees are pleased with the system as it reduces the number of lost driver records and saves a three-week wait in the filing cycle. Furthermore, the error checking routines which process the typed records provide a way for rating typists--a subject of concern to MDS personnel.

The brevity of this ex post facto analysis is indicative of the effort required for making such an evaluation of a well-planned and well-executed project. This project is one in which the chronology of Chapter 1 was followed. Specifically, the agency carefully examined its problem and formulated a project to attain specific, stated objectives. Overall components were defined and restrictions were established. In addition, sub-objectives, in this case the conversion of driver records, were established. Alternative projects were considered in this analysis, and a "best" project was selected. The project was implemented through a well-defined plan of agency responsibility and control. Progress was monitored on a monthly basis and pertinent data were recorded (number of records converted, costs incurred).

In a situation such as this extensive "after the fact" evaluation and analysis of deficiencies are unnecessary. This again points out the importance of the good agency planning and execution.

In other situations, OHSP may find that considerable ex post facto evaluation is desirable to determine why projected and attained benefits and/or costs differ. From these evaluations, the agency involved can be given information helpful in avoiding these problems in the future. In addition, OHSP can use the information to develop future programs which minimize the likelihood of such problems.

REFERENCES

1. Planek, T. W., "Total Program Planning for Highway Safety," Traffic Safety, July, 1967, pp. 8-12.

2. Planek, T. W., "Principles and Guidelines for the Development of Community Support Planning and Action, Unpublished paper, 1968.

3. Borkenstein, R. F., et al., The Role of the Drinking Driver in Traffic Accidents, Department of Police Administration, Indiana University, 1964.

4. Cochran, W. G., and Cox, G. M., Experimental Designs, J. W. Wiley, 1950.

5. Kempthorne, Oscar, The Design and Analysis of Experiments, J. W. Wiley, 1952.

6. Winer, B. J., Statistical Principles in Experimental Design, McGraw-Hill, 1962.

7. Jacobs, H. H., Conceptual and Methodological Problems in Accident Research, Behavioral Approaches to Accident Research, 1961, Association for the Aid of Crippled Children, New York 17, New York.

8. Hall, W. K., "Statistical Methodology for Highway Safety Research," unpublished memorandum, Highway Safety Research Institute, 1967.

9. Carlson, W. L., "Identifying the Problem Driver from State Driver Records," HSRI Research, Highway Safety Research Institute, The University of Michigan.

10. Cochran, W. G., Sampling Techniques, J. W. Wiley, 1963.

11. Cochran, W. G., "The Planning of Observational Studies of Human Populations," Journal Royal Statistical Association, v. 128., pp. 234-265, 1965.

12. Rapoport, Anatol, "The Alcoholic Driver: A Critique, The Prevention of Highway Injury, Highway Safety Research Institute, 1967.

13. Prest, A. R., and Turvey, R., "Cost-Benefit Analysis: A Survey," The Economic Journal, December, 1965, pp. 683-735.

14. Recht, J. W., How to do a Cost/Benefit Analysis of Motor Vehicle Accident Countermeasures, National Safety Council, 1966.

15. HEW, Motor Vehicle Injury Prevention Program, Dept. of Health, Education, and Welfare, 1966.

16. Report of the Secretary's Advisory Committee on Traffic Safety, Dept. of Health, Education, and Welfare, 1968.

17. Dawson, R. F. F., Cost of Road Accidents in Great Britain, Road Research Laboratory, 1967.

18. You and the Cost of Highway Accidents, Utah State Department of Highways, 1965.

19. Hatry, H. P., "Criteria for Evaluation in Planning State and Local Programs," in Committee of Government Operations, 1967.

20. Crumlish, Notes on the State-of-the-Art of Benefit-Cost Analysis as Related to Transportation Systems, National Bureau of Standards, 1966.

21. Webster, Lee A. and Gruen, Wayne T., A Summary of the Vehicular Speed Regulation Research Project; The Engineering Experiment Station, University of Illinois, June 1966.

22. Munden, J. M., An Experiment in Enforcing the 30 Mile/Hour Speed Limit; Road Research Laboratory, Harmondsworth, England, Report #24, 1966.

23. Shumate, Robert P., Effect of Increased Patrol on Accidents, Diversion, and Speed; The Traffic Institute, Northwestern University, Evanston, Ill., Research Project R 13, 1958.

24. Michaels, Richard M., The Effects of Enforcement on Traffic Behavior, Public Roads, Dec. 1960.

25. Maciejewski, Trooper Kenneth A., Evaluation of VASCAR; Internal Report, Michigan State Police, East Lansing Post.

26. University of North Carolina, An Evaluation of the Operational Efficiency of "VASCAR"--A Speed Measuring Device; Highway Safety Research Center, University of North Carolina, Jan. 1968.

27. State of Michigan, *Planning and Development Regions for Michigan Office of Planning Coordination*, Bureau of Planning and Program Development, Executive Office of the Governor, State of Michigan, Technical Report No. 14, Feb. 1968.