Report HSRI-71-117

PLANNING FOR CONCENTRATED IMPLEMENTATION OF HIGHWAY SAFETY COUNTERMEASURES

Volume 3: Detailed Program Plans

J. O'Day
J. S. Creswell Jr.
J. A. Green
M. E. Lee
J. C. Marsh IV
S. Schultz II

Highway Safety Research Institute
Institute of Science and Technology
The University of Michigan
Ann Arbor, Michigan 48105

August 1971

Final Report

1 July 1970 to 31 August 1971

Prepared for
National Highway Traffic Safety Administration
Department of Transportation
Washington, D.C. 20591

Prepared for the Department of Transportation, National Highway Traffic Safety Administration, under Contract FH-11-7613.

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
HSRI-71-117; Vol. 3		

4. Title and Subtitle	5. Report Date
PLANNING FOR CONCENTRATED IMPLEMENTATION OF HIGHWAY SAFETY COUNTERMEASURES Volume 3: Detailed Program Plans	August 1971 6. Performing Organization Code
7. Author(s) J. O'Day, J.S. Creswell Jr., J.A. Green, M.E. Lee, J.C. Marsh IV, S. Schultz II	8. Performing Organization Report No. HSRI-71-117; Vol. 3
9. Performing Organization Name and Address Highway Safety Research Institute The University of Michigan Ann Arbor, Michigan 48105	10. Work Unit No. 11. Contract or Grant No. FH-11-7613
12. Sponsoring Agency Name and Address National Highway Traffic Safety Administration Department of Transportation Washington, D.C. 20591	13. Type of Report and Period Covered Final Report July 1 1970 to August 31 1971 14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

In order to justify the allocation of money to existing safety programs or to allocate new funds to innovative programs, a need exists for some method of determining the efficacy of accident countermeasures.

Six program categories have been defined covering the range of activities associated with the Highway Safety Program Standards. For each of these program categories, an evaluation experiment has been designed that is intended as a vehicle to demonstrate recommended program evaluation methodology as well as a means of obtaining the required information for rating the feasible countermeasures in terms of their safety effectiveness.

Program plans for each of the six evaluation experiments are contained in this report. Experiments include the categories: 1) Road User Regulation; 2) Information Flow; 3) Road User Preparation; 4) Vehicle Regulation; 5) System Restoration; and 6) Highway Regulation.

17. Key Words	18. Distribution Statement

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	280+v	

Form DOT F 1700.7 (8-69)

PREFACE

This document is the third volume of a four volume report covering the results of a one-year study contract to determine a program planning methodology for the evaluation of highway safety countermeasures. The present volume contains detailed program plans for the six program categories defined in Volume 2. Volume 1 is an introduction and summary to the results obtained in the study; Volume 2 contains the rationale for program planning; and Volume 4 is a report bibliography on documents pertinent to countermeasure development.

TABLE OF CONTENTS

TABLE OF CONTENTS Continued

TABLE OF CONTENTS Continued

1.0 INTRODUCTION

After introduction of the sixteen Highway Safety Program Standards, a need for some positive criteria to aid in the allotment of funds for the various safety programs became evident at the national level of government. The detailed evaluation program plans presented in this document represent an attempt to satisfy this need by designing demonstration activities to measure the value of highway safety programs in terms that are understandable to the local practitioner.

In a causal chain approach to the evaluation of countermeasure effectiveness, an ordered set of steps are assumed to take place between the institution of a safety system change (or countermeasure) and the final desired effect (usually the reduction of accidents or fatalities). First, the change is assumed to modify the characteristics of some component or components of the highway traffic system. Major components are, of course, the driver, the vehicle, and the roadway. At the next stage, the change in characteristics is assumed to modify the performance of the affected part. Thus, the educated driver will execute turns in a more acceptable fashion; improved vehicle brakes will stop the care more quickly and more safely; or highway median barriers will reduce the incidence of lane crossings. Finally, at the operational stage, the modified performance obtained from an affected component is assumed to result in a new fatality rate or accident rate. In this fashion the initial safety change is linked to the final desired result (reduced accidents) by a series of steps. Furthermore, the effect of the whole process can be characterized by a number of effectivity measures at each stage of the chain.

The tendency in the past has been to define the efficacy of a countermeasure solely in terms of its ability to reduce accidents, thereby neglecting the host of other possible outcomes. Or if the intermediate effects are considered, it is often at the expense of the final outcome or of the whole range of possible effects. The real need for countermeasure evaluation programs lies in the need to obtain information on the utility of changes in the safety system that is both understandable and convincing to the local safety practitioner, i.e., the person at the local government level who is responsible for the day-to-day decisions concerning the operation of the highways. At the present time, the highway safety field is saturated with extravagant claims for every conceivable gain. It is consequently not surprising when a local safety practitioner (who is perhaps unfamiliar with the latest safety developments, but certainly not incompetent) is hesitant to accept a claim that he only needs to introduce SYSTEM CHANGE X in order to reduce fatalities by 20%.

Instead of this approach, it seems more logical to support any claims of potential accident reduction by a comprehensive picture of countermeasure effects and ramifications. While scientific formulations may be necessary in presenting this information, its basic form should be readily understood and assimilated by all the appropriate practitioners, even if they are unfamiliar with the scientific formulations.

In Volumes 1 and 2 of this report we have presented, in some detail, our understanding of the evaluation process and the program planning methodology that is needed to produce sound, usable information. This volume contains detailed program plans for the six program categories described in Volume 2. With the category descriptions of Volume 2 as background material, the plans of this volume may be related in an interconnected scheme of evaluation methodology.

2.0 ROAD USER REGULATION

The Road User Regulation program category encompasses all those safety system changes that are designed to modify the characteristics of people who utilize the highway system; this includes automobile or truck drivers, motorcyclists, bicyclists, and pedestrians. For the most part our use of the term "road user" will imply automobile drivers, but the applicable countermeasures are in no way restricted to this user category. Conceptually, there are three components of the regulation process: the regulations as exemplified by the system of codes and laws and their interpreters, the courts; the regulators or police that are commissioned to enforce the regulations; and the regulated, the objects of the regulation process.

Because Road User Regulation program acts on people, the real effect of any given accident reduction technique is notoriously difficult to assess. In practice there are many operational program personnel who operate much of the time on the basis of hearsay evidence or anectodal accounts of successful methods that have been used elsewhere. Thus, equipment is purchased and new enforcement procedures are put into operation on the evidence that a similar method was used at another site and was successful there (at least in the personal opinion of some person who was involved in that effort). Very few meaningful studies that have statistical validity have been conducted to determine the efficacy of enforcement measures even though there is a strong feeling that they must be good. One of the most well known is the Operation 101 work done in California. This study has developed experimental evidence of a relationship between enforcement and accident reduction, but it was not able to define the quantitative nature of that relationship. A more recent study called Operation 500 is designed to supply this relationship.

Partly because of this lack of adequate knowledge, the city councils and state legislatures who must appropriate the necessary money for funding any proposed program have, in large part, remained unconvinced of the value of specific people regulation countermeasures. A prime example of this occurred in the city of Flint, Michigan. Coincident with the introduction of a selective enforcement program, the city enjoyed a reduction in fatalities from 31 in the year before enforcement to 13 in the year of the experiment. The operating agencies would like to give the entire credit for this change to the enforcement program, and at a currently accepted value of a life lost in traffic of $140,000, the police program should be worth $2,500,000 to the community. Program cost, on the other hand, was only $400,000, so that the cost benefit of this program is apparently large. However, in spite of this evidence, the program administrators are having rough sledding in their attempt to convince the city council to continue the activity.

Consequently, if there are truly a number of road user regulation countermeasures that are worth more than they cost (in some sense), then the demonstration efforts must be careful to collect comprehensive proof of this in a way that will serve to convince the local decision makers.

2.1. A Review of Regulation Activities

The Road User Regulation program area interfaces in a fundamental fashion with four of the sixteen Highway Safety Program Standards. They are the Codes & Laws, Traffic Courts, Alcohol Safety, and Police Traffic Services standards. Because the alcoholic driver is under intensive study in the ASAP programs sponsored by the NHTSA, the Alcohol Safety Standard has received minimal attention in this study except where it interfaces with the remaining regulation activities in the large scale evaluation programs. In our assessment of program priorities (as discussed in Volume 2 of this report) the Police Traffic Services Standard has received primary consideration since it appears that this area is most likely to be actively pursued as a realistic endeavor in the near future.

The principal aim of the Codes & Laws Standard is to achieve uniformity among the various laws and ordinances set forth by the states to attain safe use of the highway system. The Traffic Courts Standard, on the other hand, seeks to achieve prompt and impartial adjudication of cases by specifying the necessity for court settlement of all moving traffic violations and by the modification of court operating procedures. Neither standard has received much in the way of "402" program support from the states. HSRI studies indicate that the state spending the largest percentage of its funds in these standard areas for the period 1967-1970 was Mississippi. There, less than 6% of the available funds were spent in the Codes and Laws area and less than 10% in the Traffic Courts area. (See Appendix A in Volume 2 of this report for a discussion of the HSRI "402" project file and its use in this research program.)

In spite of this apparent lack of concern in the "402" area, there is a great deal of concern in the society at large that the court system is falling far short of its goal of providing rapid, just treatment of the cases that it treats.

The prime responsibility of the police force in safety activities is to insure that the codes and laws that have been established for driver behavior are adhered to, and that those persons apparently deviating from the established norms of safety and legality are apprehended and referred to the court system for ajudication and, if convicted, correction. The standard seeks to upgrade the effectiveness of the police in carrying out this responsibility by providing training, developing better reporting techniques, and promoting better management of the personnel available to the police force.

In an extensive survey of the Police Traffic Services (PTS) literature (Reference 2-1), police services have been divided into three primary areas:

1) Street Functions - Traffic control, enforcement, accident management, motorist services, etc.

2) Support Functions - Court appearances, education activities, accident analysis, etc.

3) Management Functions - Financial control, manpower deployment, plant maintenance, staff training, etc.

The PTS field is a popular one for state and local programs since 31% of the "402" projects and 18% of the funding has been expended in this standard.

Because the PTS area involves a great variety of tasks, there are a bewildering number of possible countermeasures that could be used in any particular situation. Hence, for the purposes of this report, we feel that it is not possible, or even, desirable to select a set of countermeasures for a particular site. The choice of appropriate change programs should be based on analyses of local needs. We do, on the other hand, discuss evaluation methodology, and we use specific programs as examples.

From our review of the literature and the programs that have been carried on in the Road User Regulation category, the applicable countermeasures have been segreated into eight generic groups. In deriving these groups, a scale of interaction between the driver and the regulation system was used for categorization purposes. That is, the groups range from internal management functions of the police force that require little contact with the drivers who are the subject of the program, to the treatment of social-individual problems that are primarily concerned with a direct contact. The countermeasure groups are:

1) Management improvements are overlaid on an ongoing PTS program where techniques like manpower allocation and cost/benefit methods are implemented.

2) Training programs where present PTS manpower is trained for some selected advanced skills (behavioral science, management, etc.) or in the use of some specific tool (eg. radar, etc.).

3) Manpower additions where more manpower, either at or above the current training level, is added to a PTS force.

4) Equipment additions where major investments in electronic enforcement gear are made along with some minimal training effort and implementation scheme.
5) Procedural clarifications in Codes & Laws where streamlining, standardizing, and training occur dealing mainly with the optimized usage and administration of old laws as well as the implementation of new laws.
6) Logistical improvements in the handling of court related problems where procedures and communications between PTS units and courts are made more efficient.
7) Media efforts where increased communication between the public and the police agencies are attempted.
8) Social-Individual problem handling techniques are implemented where a system of direct contact actions is used to influence problem drivers such as alcoholics.

The use of a program matrix to determine the additive and interactive effects would be prohibitively expensive and time consuming. For instance, the combinations of eight interactive countermeasures groupings at three intensity levels for each group would require 747 experimental determinations. Carrying out such a program would be clearly ridiculous. Our proposal, therefore, is to use the master "building block" program plan discussed in Volume 2 as a long term planning guidance approach to countermeasure evaluation.

2.2 The Program Design Rationale

Due to the changing population patterns in our country, it appears that cities in the range of 100,000 to 200,000 residents are experiencing the greatest rate of increase of highway accidents. Several reasons for this, at least on the surface, can be advanced. Large cities in the past decade have, for the most part, been suffering a population decline as a result of migration to the suburbs and to smaller cities. Thus, the facilities that have been successful in controlling accidents in the past are certainly adequate to handle the needs of a reduced population if adequate funding is available. Small cities and rural areas, on the other hand, have not grown to the point where accidents have become a dominant problem. In the medium sized cities of the size under consideration, however, the rapidly expanding population has in most cases outstripped the facilities available so that real problems exist in the accident control field (as well as in schools, crime, etc.).

As a result of this currently important factor, we have chosen to select programs that will produce results that can be used by these cities to allocate their meager and usually hard-fought-over facilities to the best advantage. Consider, then, the application of people regulation countermeasures to a city in the one hundred to two hundred thousand population range. A further breakdown of the program implementation is considered important here. It appears that two types of experiments could profitably be conducted to answer immediate highway safety questions in these communities as well as to indicate what measures may be optimal in the long run.

The determination of short range needs is attacked by the development of program plans that could be implemented in any typical community with defined problems using measurement techniques and instruments that can be conveniently added to the site. We have elected to show how this may be accomplished by discussing the current plans for a Selective Traffic Enforcement Program (STEP) in Chattanooga, Tennessee. The techniques and countermeasures to be defined for this community have been drawn from that city's proposal submitted to NHSTA (Reference 2-2). We believe that implementation of additional programs of this general type could be accomplished within the "402" program framework, with modifications in usual operating procedures to enchance the recording of suggested measures of effectiveness. The experimental programs conducted in this manner would upgrade the capability to identify successful programs, and publication of this information would prompt other jurisdictions to use new methods.

Selection of the Chattanooga proposal for detailed discussion was motivated by two concerns: 1) that the countermeasure groups be representative of possible evaluation methodologies, and 2) that a concrete example of likely evaluation problems be kept firmly in mind. The four groups used in Chattanooga (media, manpower, training, and social handling) represent the most commonly applied countermeasures. Moreover, the methods used in evaluating these four groups are quite analogous to the techniques necessary for evaluating other groups: for example, management and equipment most often are directed at enhancing police productivity in apprehending violators, as is the training function. While we have recommended a rather full set of techniques, each of them requires only a modest effort which in most cases can be performed by local personnel. Implementation of all the techniques will cost more than is currently budgeted for evaluation in the Chattanooga plan, but a full effort will still not be overly costly considering the possible national significance of results and the long-term costs of operations if they are to be continued.

At the other end of the evaluation spectrum, there is the need to determine the optimal allocation of funds to people regulation projects. To satisfy this need for information, we suggest the selection of several cities in the 100,000 to 200,000 population range as "laboratories" for the ongoing evaluation of accident reduction countermeasures. These cities would receive a heavy investment in equipment (radars, buried loop detectors, centralized computer data gathering and control center, etc.) that would permit the rapid and thorough evaluation of many sophisticated techniques. A basis for this approach is exemplified in the city of Bloomington, Indiana. Here, a continuing set of instrumentation has been designed to unobtrusively sample traffic activity throughout Monroe County. This instrumentation has been developed to a relatively high state by personnel at Indiana University, providing a very useful test base (Reference 2-3).

Such laboratories will serve a continuing national purpose by providing other communities with soundly based information on which to base their programs. Countermeasure programs will be selected for testing and implementation by joint agreement between local officials and a national advisory board. The constitution and activities of this board will be discussed in a subsequent section.

Few communities remain stagnant with regard to changes in the road user regulation area, but changes ordinarily do not come very rapidly, either. A new traffic enforcement division may be created, a new police chief with a particular interest in traffic safety may be hired, a judge may institute a new offender treatment program, or a legislature may pass a new rule of the road. Yet, for the most part such changes come one at a time, and seldom are they made according to a long-term plan or are their effects carefully evaluated.

In contrast, the national laboratories will allow a planned and relatively rapid introduction of innovations into the road user regulation activities of a community. These innovations will be carefully measured, and their effects well evaluated. The subsequent widespread publication of the results will then serve to guide other communities in their efforts.

2.3 Specific Programs

Four programs have been selected for detailed analysis: they are public information, police training, increase enforcement activity, and a post-conviction defensive driving course. These programs are based on a proposal made by Chattanooga, Tennessee under the Selective Traffic Enforcement Program (STEP) of the National Highway Traffic Safety Administration (NHTSA). As indicated previously, these programs represent only one of many combinations that can be attempted in the road user regulation area, but this program seems quite typical of those envisioned for many areas.

The long-term goal of all these activities is to reduce the frequency of traffic accidents in the area in which they are applied. It is doubtful that attainment of this objective can be measured considering the wide variety of other accident producing factors over which the program has little control. Nevertheless, detailed information must be maintained on traffic volumes and composition, population changes, exposure measurements, and accident information to attempt to see if any change has indeed occurred.

The intermediate goal of the activities is to reduce the incidence of hazardous actions by drivers in the area of concern. It is proposed that a number of measurements be conducted to observe these hazardous actions at selected points in the community and to see whether significant changes occur after appropriate compensation for changes in traffic volumes and densities, local regulations, and engineering changes. More detail will be given on the nature of these measurements in the section on increased enforcement that follows.

Immediate goals can only be discussed in terms of the specific activities considered. For example the immediate goal of the public information activity is to alert some aduience to the possibility of increased enforcement action, and an immediate goal of manpower training is to increase the capability of officers to present cases in court. Measurement in terms of these goals will serve to identify effects of particular programs; if a specific program does not have an effect on the immediate goal of that effort, it can be stated with near certainty that it has not affected the intermediate and long-term goals of the entire project. If all or more than one of the efforts attains a reasonable amount of performance on the immediate goals, then it will be impossible to determine which effort made the greatest relative contribution to the overall effort. Suppose that both the training program and the increased enforcement presence program attained some degree of success;

then it will be impossible to tell if a concommitant drop in the violation rate is due to the fact that the better trained officers were more proficient in apprehending and prosecuting offenders, or if the drop results simply because more policement are present in a specific area.

2.3.1 Public Information

The public information campaign associated with the program seeks to make the public more aware of the activities of police and courts in dealing with hazardous traffic actions. In increasing this awareness two closely related objectives may be sought: (1) drivers can be motivated to perform their driving tasks better, even without a concern about potential sanctions; and (2) even lacking a desire to drive better, drivers may control their behavior more carefully simply out of concern with possible legal sanctions. Both effects will be present in all drivers, but to a varying degree in any one driver. The ideal campaign can be directed to emphasize both motivations with a message simultaneously promoting good driving practice and warning that if such good practice does not occur, the local police and courts are now prepared, more than ever, to deal with poor practice.

The message can be presented through a variety of media: e.g., newspapers, television, radio, posters and billboards, bumper stickers, addresses to civic and school groups, and simply by word of mouth. The content of the message can be varied from source to source depending on the audience that it is intended to reach. Local information on the readership, viewers, or listeners of various media can help guide the presentation. For example, the positive aspects of better driving skill can be aimed at newspaper readers, television audiences, and civic club meetings, while a more mixed message about both the positive and negative reinforcements present in the program can be directed through radio and school meetings. The word-of-mouth approach cannot be overlooked. If the police in the course of their normal contacts with such groups as drive-in restaurant patrons, operators of "speed" shops, bartenders, and insurance agents dealing in assigned risk policies spread the word that enforcement will be getting much tighter, a large group of people not normally in great contact with conventional media can become aware of the program; this group can possibly also contain a greater proportion of high violators than is found in the general population.

The best means for directly measuring the effectiveness of publicity is to interview persons who may be exposed to the publicity. This interviewing process can be conducted in a number of ways. The most expensive, but also the most precise, way is to conduct a household survey of persons in their homes. The second most expensive way is by telephone interviews. A poll of people in public places is a low cost means of obtaining information, but in this method the sample can be greatly biased by the poor representation of general characteristics shown by groups assembled in any one place and by unconsious or conscious biases of the interviewer in picking particular persons to interview.

For the present purpose, telephone interviews seem to be most desirable since a large number of persons can be contacted in a short period of time at a relatively low cost. Only a few short questions need be asked to determine if people have heard the message, have understood the message, and believe the message. One such approach is to build the information campaign around some slogan such as "Keep in step with STEP". People can then be asked if they have heard it or seen it. If they have heard it, they then can be questioned as to where they heard it and to explain what it means. Responses can be graded from 1 to 5 depending on the completeness and accuracy of the answer, 5 being they know it means "Selective Traffic Enforcement Program", and 1 being a totally incorrect answer such as "A new car advertisement". Regarding the believability of the advertisement, one can ask whether people feel that traffic law enforcement in the area will be "Much greater", "Somewhat greater", "About the same", "Somewhat reduced", or "Greatly reduced" as a result of the program.

Responses to such questions can be related to factors such as age, sex, and media source to determine which areas are making the greatest contribution to the program. The interview process can be conducted both before the campaign commences and after the campaign has been in progress for some time. The immediate measure of success will be if the number of people having heard the message, understood it, and believed it increased significantly after the campaign. Some attempt may be made to determine whether people intend to change their behavior as a result of being exposed to the campaign. A sample interview is presented in Figure 2-1.

A sample size of 500 interviews conducted once before and once after the campaign will be adequate to determine with good confidence a ten to fifteen percent change in audience response. The sample can be obtained quite easily with the use of eight to ten phone interviewers for a period of a week. The calls will be placed from six to ten p.m. Names of persons to be called will be chosen at random from the alphabetical listings. A minimum of two calls for two nights will be made in an attempt to contact individuals who do not answer the initial call. A certain amount of persuasive effort will have to be used to convince those who are initially reluctant to cooperate in order to minimize non-response bias. Arrangements will have to be made to provide the listed phone number of a recognized city agency, such as the police department, for those who wish to call and verify the legitimacy of the survey.

In summary, the goal of the public information effort is to make people aware that an enhanced enforcement program will be in effect. The purpose of this effort is to get people to alter their driving to avoid hazardous actions. This effect

FIGURE 2-1. SAMPLE PHONE INTERVIEW

Hello, I'm (interviewer code name), and I am conducting a survey for (city agency). It will take just a few minutes of your time. Your answers are very important to us in our present highway safety project.

(1) Are you a licensed driver? Yes No
(If the answer is "no", ask if a licensed driver lives in that household. If a licensed driver does live in the household ask to speak with him or her and explain that the survey is directed toward drivers. If no licensed driver resides in the household, again explain that the survey is for licensed drivers apologize for the inconvenience, and terminate the call.)

(2) Your age? _____

(3) Your sex? _____

(4) Your occupation? ______________________________

(5) How many miles do you drive in a year? _____________

(6) Have you heard the phrase "Keep in step with STEP?" Yes No
(If the answer is no respond by asking if they recall hearing any traffic safety message lately. Record the phrase here ____________________. If the response resembles the program phrase, continue the interview as instructed. If there is no resemblance to the phrase, ask where they heard it and record this information here________________. Then terminate the interview.)

(7) Can you explain what it means? ______________
(Award points on scale:
5 Selective Traffic Enforcement Program
4 Police Department Traffic Project
3 Some type of Police Department project
2 Some type of city program
1 Inappropriate response.)

(8) Where did you hear this phrase?

___Newspapers
___Television
___Radio
___Sign
___Meeting
___Bumper sticker
___Friends
___Other(record)__________

(List the order in which they give responses such as radio, meeting, sign as 1,2,3 on appropriate line.)

(9A) Do you think that in the near future there will be any change in the way the (city name) Police Department will be enforcing traffic laws? Yes No
(IF THE ANSWER IS YES ASK:)

(9B) How much of a change do you think there will be? _______
(Grade the response as:
1. A very great increase in enforcement.
2. An increase in enforcement.
3. A decrease in enforcement.
4. A very great decrease in enforcement.)

(10A) Recently, have you notice any change in how well your friends observe traffic laws? Yes No
(IF YES ASK:)

(10B) What kind of a change have you noticed? _______
(Grade the response as:
1. Very much more carefully.
2. Somewhat more carefully.
3. Somewhat less carefully.
4. Very much less carefully.)

Thank you very much for your time. Your answers have been very useful to us. Good night.

cannot be directly measured, but the immediate effect of the program can be measured through a survey to determine if specific audiences heard, understood, and believed the information concerning enchanced enforcement.

2.3.2 Police Training

The police training course is outlined in the Chattanooga proposal. The program emphasizes improving the quality of the enforcement activity through better knowledge of enforcement techniques and of legal procedures. To a great extent these items can be measured objectively and directly in terms of enhanced performance of police duties. A first measurement can be obtained by giving the officers a test of their knowledge of the elements contained in the course both before and after the training is completed. This, at a minimum, will tell whether they have learned the knowledge which the course sought to convey. This cannot be the only measurement though, since it does not test whether the knowledge conveyed was relevant to the problems which the training sought to solve.

Two major purposes of the training activity are to improve the quantity and the quality of the enforcement effort. The quantity can be measured simply in terms of the number of enforcement contacts made per man-hour assigned to traffic patrol before and after the training. The quality is somewhat more difficult to measure and even to define. One way of looking at the quality of the contact is to examine whether the contact served to alter the violator's future performance; but this is difficult to assess and is affected by a number of variables not under the control of the experiment. What one can do is to define quality as some parameter of the seriousness of the offense which the officer detects, apprehends, and prosecutes successfully. One can weight the arrests by the point scale used in determining driver license suspensions. This measure will be adjusted to some extent for differences in duty parameters such as shift that the officer's work and the character of the patrol area.

An officer working in an area with many bars on the 11:00 p.m. to 7:00 a.m. shift will have many more opportunities to make arrests for driving while intoxicated. Conversely, the officer working in an upper income residential area during the 7:00 a.m. to 3:00 p.m. shift will have to be assessed against performance in that district on that shift. Table 2-1 presents some performance measures to be used as intermediate criteria for the training.

TABLE 2-1. POSSIBLE MEASURES OF EFFECTIVENESS OF TRAINING

Raw Measures	Weighted Measures
Number of violations detected per patrol hour.	Index of violations detected per patrol hour.
Number of violators contacted per patrol hour.	Index of violators contacted per patrol hour.
Number of violators arrested or issued summons per patrol hour.	Index of violators arrested or issued summons per patrol hour.
Number of convictions on original charge per arrest.	Index of convictions on original charge per arrest.
Number of convictions on reduced charge per arrest.	Index of convictions on reduced charge per arrest.
Each measure to be constructed by officer, shift, and district for six months before training and six months after training.	Each index to be constructed by officer, shift, and district for six months before training and six months after training.

Improving the officer's ability to present cases in court through better knowledge of the legal requirements and through improved courtroom demeanor can be measured in two ways. The effectiveness measures based on conviction experience can be one form of this measurement, but since most traffic offenses result in non-contested guilty pleas , the statistics may not indicate too much increase in performance. Another approach is to interview other court participants to see if they feel that the officer's performance has been improved following the training. This interview should be conducted at least two months before training and at least three months after training and should involve traffic court judges, prosecuting attorneys, defense attorneys, and sundry court personnel such as clerks and bailiffs. For each of 100 cases before and after the training, the individuals will be asked to complete a questionnaire. The questionnaires will be evenly distributed over the number of officers trained so that the performance of any one officer will not greatly affect the results. A sample of the questionnaire is presented in Figure 2-2.

Assessment of the value of training in accident investigation is the most difficult part of evaluating the training program. Short of having a team of expert accident investigators reinvestigate the accident, only some limited measures can be made of the improved investigatory skills. An observer can review the accident reports completed by the officers before and after training for such elements as neatness, completeness, and attention to detail and examine the report for internal consistency (such as recording the vehicle's estimated speed as in excess of the posted limit, but not recording this as a violation associated with the crash). One can further assess the quality of the training by looking at the arrest and conviction rates for offenses detected in accidents, although this measure can be distorted by changes in departmental enforcement policy. What one wants to test here is the hypothesis that given a fixed department enforcement policy and a constant amount of supervisory prompting, the better trained officer will discover and prosecute successfully more offenses arising out of accidents than an untrained officer. Changing either the department policy or the degree of supervision will mask the effects of training.

The intermediate goal of training is to contribute to the reduction in hazardous acts, but as in public information campaign it does not seem possible to measure directly the effect on the commission of the undesired actions. What can be measured though is the effect of the training on the productivity of the officers in enforcing the laws. This productivity

FIGURE 2-2. SAMPLE QUESTIONNAIRE: COURTROOM PRESENTATION ASSESSMENT

The purpose of this questionnaire is to assess the effectiveness of (city name) police officers in presenting cases in court. All answers will be held strictly confidential. Data is being gathered for statistical purposes only and will not reflect on the performance of any individual officer. Please do not give any indication of your name, officer's name, or defendant's name.

Date:____________________

Original Charge:________________________

Disposition: (Please circle appropriate answer)

1. Defendant pleaded guilty to original charge.
2. Defendant pleaded guilty to reduced charge of _______.
3. Defendant pleaded innocent and was convicted by jury trial of original charge.
4. Defendant pleaded innocent and was convicted by jury of reduced charge of ______________.
5. Defendant pleaded innocent and was convicted by judge of original charge.
6. Defendant pleaded innocent and was convicted by judge of reduced charge of ______________.
7. Defendant pleaded innocent and was acquitted by jury.
8. Defendant pleaded innocent and was acquitted by judge.
9. Defendant successfully challenged validity of charge, arrest, or evidence and case was dismissed.

How would you rate the overall conduct of the officer in presenting his case? (circle answer)

1. Excellent
2. Good
3. Average
4. Unsatisfactory
5. Poor

Would you say that the officer's conduct was (circle answer)

1. Very professional and well mannered.
2. Normally professional.
3. Indifferent.
4. Somewhat unprofessional or discourteous.
5. Hostile and bordering on contempt.

In presenting evidence was the officer? (circle answer)

1. Very clear, concise, and logical.
2. Reasonably understandable.
3. Somewhat perfunctory.
4. Rather vague and rambling.
5. Completely confused and unclear.

Would you say that the officer appeared to act toward the defendant (circle answer)

1. In a fair and courteous manner?
2. In a neutral manner?
3. In an indifferent manner?
4. In a discourteous manner?
5. In a rude and hostile manner?

Would you say that the officer's knowledge of the law and rules of evidence in presenting his case was (circle answer)

1. Excellent?
2. Good?
3. Average?
4. Unsatisfactory?
5. Poor?

Was the defendant represented by counsel? Yes No (circle answer)

If the defendant was acquitted or the charge dismissed, what was or what appeared to be the principal problem with the case?

1. Strong factual situation but poorly presented by officer
2. Weak factual situation but strongly presented by officer
3. Both weak factual situation and poorly presented
4. Inappropriate application of law, i.e. faulty charge
5. Faulty arrest procedure or improperly gained evidence

What is your relationship to the case? (circle answer)

1. Judge
2. Clerk or bailiff
3. Prosecuting attorney
4. Defense attorney

is measured directly through the statistics proposed and indirectly through observation of the officers' effectiveness in courtroom presentation and in accident investigation.

2.3.3 Increased Enforcement Activity

Numerous attempts have been made in the past to measure the effect that increased enforcement activity has on such matters as violation rates and accident rates. Such studies have been reviewed by several authors with mixed results (see References 2-1 and 2-3). Generally one concludes that there is little undisputed success, principally because of contamination of the experiments with unknown factors, or because of the lack of controls. Similarly, of course, there is not any evidence of failure. With appropriate cautions as to the difficulty of attaining valid results, one can still try to form an experiment.

The number and type of violations committed on a particular segment of highway depend on a large number of factors. A nonexhaustive list of these factors includes the characteristics of the drivers on the road, the time of day, day of week, weather, light, time of year, holiday or workday, the traffic density, and the structure of local codes and ordinances. Given all these factors, one may attempt to vary the intensity of enforcement in terms of the number of police enforcing the law, and the strictness of the enforcement policy. If all goes well and none of the other parameters change radically, the hypothesis is that as the intensity of enforcement increases, the frequency of violations (either detected or undetected) will decrease.

Changes in Violation Frequencies

Two approaches can be used to measure the change in violation frequencies. The first is to compare the frequency of violations in a particular area before a program is instituted, while it is in effect, and after it is completed. One can use several techniques to determine the violation frequency. The before, during, and after technique has several attractions in that it minimizes the effect of environmental differences by hopefully using the same environment, i.e. road parameters, and population during the course of the entire experiment so that observed effects can be attributed to the program. However, some care must be exercised in this technique. First, it cannot be simply assumed that the environment remains constant or that the population remains the same. A careful check must be maintained on both changes in the physical characteristics of the road such as construction and traffic engineering and in the type of persons using the road. A most useful technique is to conduct an origin-destination study in each of the time periods to ascertain what changes, if any, have occurred in the road use. If there are substantial changes in the mix of people using the road, say from younger drivers on recreational trips to older drivers on communting trips, the frequency of violations can drop quite independently of the enforcement efforts. Second, one must check for time trends in violation activities which may fluctuate in a purely random manner. For example, the period immediately prior to the instituting of an enforcement campaign may be one of unusually high violation action. This will be partricularly true if the area is selected for attention because it has recently become a problem area, either because of an unfavorable accident experience or because of numerous citizen and police observations of a high incidence of traffic violations in the area. If this simply is due to the chance occurrence of an unusually poor period, the period of intensive enforcement will almost certainly be one of lower violation experience even though the lower violation experience has not been caused by the enhanced enforcement. (See for example, regression-to-the-mean discussion in Volume I of this report.) Conversely, violations by chance can be low during the time immediately preceeding the experiment, and consequently, the observed effects of the experiment can be masked by a return to normal from an unusually slack period. This is even further complicated by seasonal variations if the pre-experimental data gathering is conducted for a period of less than a year. Unfortunately, many demonstration programs are not so structured so that these seasonal and cyclic factors in violation activity can be carefully observed. The only resort then is to conduct pre-institution measures for as long as possible before actual operations must begin and then collect data during the experiment for a long enough period that the effects of such factors can be observed or averaged out. A minimum of two years will be necessary for these types of observations and three years will be preferable.

The second means of measuring the effect of the enforcement activity is to find locations which in some way match those that are being subjected to the enhanced enforcement program. This is not an easy task for no two areas are exactly alike. Operationally the best that can be done is to find roads of a similar type, i.e., urban collector streets or rural two-lane highways, with similar traffic volumes serving areas with similar demographic characteristics. The use of Census data tied to specific streets as proposed in Chattanooga will be excellent for this matching purpose. A further criterion, which may be difficult to meet in a local demonstration program, is that the control street be in an area remote from the test street to minimize spill-over effects from the test to the control area. As with the test street, observations will be conducted on the traffic volume, traffic make-up, and road characteristics of the control street on a before, during, and after basis. If these parameters do not vary markedly from the control to the test area over time, one can hopefully isolate the effects of the increased enforcement activity.

In the context of the Chattanooga program it is suggested that observations be conducted at two sites along each of the three road sections designated for increased enforcement, i.e., McCallie Avenue, Bailey Avenue, and the I-75 freeway. Similar areas should be selected in other parts of the city as control sites. The twelve control sites should be matched as closely as possible with the twelve test sites. Observations of violations and hazardous acts should be made on a systematic basis among all the sites. These observations should be structured over time so as to cover at each site all combinations of time of day (divided into four, six-hour periods, or eight, three-hour periods), day of the week, and season of the year. This will yield a total of 112 to 224 observations per site per year. Observations should be varied among the sites so that no site has data collecting activities clustered at one point in time. Use of the Latin square design is suggested.

Observations will be conducted in an unobtrusive manner so that neither police nor motorists will be acutely aware of being seen. The reason that the motorists should not be aware of the observations is that, having been sensitized by publicity to the STEP program, any obvious recording of their driving behavior may affect their usual responses to particular traffic situations. They might reasonably be acting under the assumption that the observation was a direct part of the enforcement effort and consequently they might make special effort to avoid committing violations. Similarly the patrolling police officers should be kept relatively ignorant of the proposed schedule, and their behavior immediately prior to and during an observation should be closely monitored. This is to inhibit the natural tendency to enforce particularly at the time and place of a scheduled data taking operation. No insult is intended by this comment, but it is a natural human impulse, particularly among those who are well motivated and trying to do a good job, to try to appear in the most favorable light. In other words, the police department will have a good program, and they will want people to know it, which will lead to the quite natural tendency to try particularly hard when they are being observed.

A number of techniques have been examined under NHTSA contract for sampling violations in the vehicle stream. They range in complexity from rather sophisticated combinations of time-lapse photography and digital recording to the most traditional method of positioning an observer with manual counters. In the Chattanooga experiment, two techniques are recommenced: use of buried loop detectors attached to a portable digital recorder and human observers. Needless to say, the observation techniques must be tied to the behavior that one wants to measure. In the present case, the principal behaviors that the program tries to control are deviations from some desired speed in the traffic stream, and those violations which often lead to conflicts and accidents among vehicles, such as improper turns, weaving in and out of lanes, and improper observance of traffic signals. A combination of human observers and the use of loop detectors seems to be best suited for these purposes.

Loop detectors have been long used in traffic engineering work for the measurement of traffic parameters such as vehicle counts. Installation of these detectors can be accomplished at a moderate cost which can be easily amortized over the period of a two or three year demonstration program. The loop detectors will be attached to a central control box at each site. There will be two loops per traffic lane to give the appropriate speed information. The central control box will contain the necessary electronic connectors and interfaces so that a portable digital recording package can be plugged in at the site in a short time. Data will be recorded at the site for the designated interval. The recorder will then be collected along with the magnetic tap containing the data. The magnetic tape will be sent to a designated computer facility for transfer and analysis by the evaluation group. The recorder with a new tape will then be moved to another sampling location and the process repeated. The use of the loop detectors is recommended over the use of pneumatic hose or doppler radar detectors for measuring spot speeds at the sampling locations. The pneumatic hose technique requires additional effort to install the hoses which can become burdensome over a long period of time, and the presence of the hoses in the pavement reveals some measurement is being conducted at that site. Further the pneumatic hose device has been used in the past in many areas as a speed detection device for law enforcement, and many motorists are sensitized to the appearance of these hoses in the roadway. Use of the radar is also ruled out on the grounds that the radar does not provide a great deal of information on the lane distribution of traffic or of differential speeds of traffic. Further, unless the radar is cleverly concealed, the tell-tale antenna apparatus is likely also be be interpreted by the motorist as a speed enforcement operation with the concommitant changes in behavior.

A slightly different strategy than the use of the portable recorder at the loop sites can be used. Instead of moving the recorder to each new site according to the experimental plan, telephones and associated data sets can be installed; when it is time to sample at the particular site, the telephone will be dailed, and the data will be transmitted to the central recorder. The principal advantage of this system is that it greatly reduces the expense of moving the recorder and the necessary time gap in observations during the move. While no formal estimates have been made, it does not seem unreasonable that the cost of the loop-to-phone interfaces will be the same as the loop-to-recorder interfaces. The cost of the twelve phones and data sets must be balanced against the saving in labor in moving the recorder and the convenience and security of a central site.

The amount of data gathered from each site will depend on the traffic volume at that site. For a location with an average daily traffic count of 6,000 vehicles, it is extremely unlikely that fewer than 100 vehicles will pass during even the slowest six-hour period. With this sample size, extremely precise estimates can be made of the mean speed, the variance in speed, the proportion of vehicles exceeding the posted speed, and the 85th percentile speed. Consequently, even very small changes in speed parameters can be detected with high confidence thus yielding a sound basis for evaluation.

Human observers will be used to record non-speed violations and other hazardous actions which are of interest to the project. The purpose of recording the non-speed violations is to cover for evaluation a more complete set of intermediate goals than measured by speed reduction alone. Recording other hazardous actions serves to push the evaluation one step closer to a direct measure of accident change' it will also be useful to the local authorities in pinpointing conditions requiring attention. Perkins and Harris of General Motors Research have developed the technique of conflict measurement; for example, one such conflict is a vehicle braking or changing lanes to avoid a turning vehicle. This action represents a perturbation in the traffic stream which could lead to an accident. The developers have demonstrated that the number of conflicts at a particular intersection is closely related to the accident experience. Consequently, reduction in conflicts can be used as a proxy for the achievement of the ultimate, but probably unmeasurable, goal of accident reduction.

The observers will be given data forms and mechanical counters on which to record information on traffic flow, violations, and conflicts. Depending on the characteristics of the location and the number of actions to be recorded, one or two observers will be assigned to each observation period. As with the speed recording data, two locations are recommended for each test and control street. Due to the greater expense of collecting this information, a less detailed sampling plan will be used. Each site will be observed a total of four times every six months. The four visits will cover day (6 a.m. to 6 p.m.) and night (6 p.m. to 6 a.m.) conditions for weekdays (6 a.m. Monday to 6 p.m. Friday) and weekends (6 p.m. Friday to 6 a.m.

Monday). Assuming that the day shifts will require two men and the night shifts one man observing at one time, the total effort for a year's data collection for twelve locations will be 1728 man-hours. Data collected will be reduced for computer processing and included in the evaluation. Sample sizes cannot be predicted without knowing the characteristics of each location, but it is anticipated that they will be large enough to guarantee good statistical precision.

Two principal problems arise with human observers: reliability and obtrusiveness. The issue of reliability centers on three elements. First, the recording of traffic data is in general a boring, repetitive, and uninteresting task. Persons who are not properly motivated will soon become quite careless in recording the data which can easily destroy any experimental effectiveness. Second, occasionally observers can have a high interest in the outcome of the experiment which can lead them to alter data to reach the desired outcome. Third, some individuals can feel that their performance is judged by the number of violations or other items they observe, and thus they over report events. To avoid these problems the project should use individuals who appreciate the need for accurate scientific measurement and the assignment periods should be relatively short so that interest is maintained. Use of part-time college students in the physical sciences or social sciences who have had an academic emphasis on data analysis is helpful. These individuals are likely to have an appreciation of the need for accuracy, and since the project is not associated with their principal concerns in life they are less likely to be subject to boredom and to other morale problems that full-time permanent personnel can experience. However, close checks must be kept on any observer's performance with such devices as supervisor-conducted control counts at the same locations as the personnel are working and/or analysis of counts taken at the same location by other workers. While some fluctuation is to be expected and a long-run decline in events is anticipated, sharp fluctuations must be questioned carefully. Frequent field checks are recommended. An observer who consistently fails to produce results which are roughly comparable to other observers (such as extreme undercounts) should be shifted from the job.

The human observers will also be somewhat obtrusive. Motorists will be able to see that they are being counted in some manner, and this may affect the motorists behavior. The human data collectors should be in as inconspicuous a location as possible so that most motorists passing by the observation point will be unaware that they are being observed until they have a chance to display the behavior under consideration. It will be impossible to avoid having the police know that the observers are present at a particular location at a certain time. If for no other reason than the personal safety of the observers, the police must be notified that an observation is being taken so that they can check to see that all is well on a periodic basis. Even if no advance notification were given of the observers' presence, doubtless the alert patrol officer will soon discover that the individuals are operating in his area. There will be obvious chances for the police to affect the outcome of counts, but they can be offset by firm command control of the individual officer's behavior. Hopefully also the police will be occupied enough with their regular duties that they simply will not have time to affect the results of the observations, particularly if these observations are carried out over an extended period of time in numerous parts of the city.

The use of the closed circuit television camera as suggested in the Chattanooga proposal does not seem to be a particularly useful technique. The disadvantages of the system are that it is quite expensive in initial capital expenditure and requires an extensive amount of effort involving the use of heavy equipment to transfer the recording equipment from one location to another if the camera is to be mounted at any height above the point to be observed. Presumably the output of the camera must be viewed either remotely at the time the observation is made or it must be recorded to video tape and viewed later by a person to record data. In either instance, the same amount of manpower is required to conduct the actual observation and data taking as is needed for the human observers, and there is the additional installation expense which is not encountered with the human observer. The one advantage that the closed circuit camera has over live human observers is that the information can be reviewed and more accuracy may be gained in the data recording process. Nevertheless, the cost and other difficulties of the television operation are enough to make its use inadvisable.

The principal effort of the project is to increase the quantity of enforcement in specified areas. The desired effect of this increased enforcement, and of the other activities in the project, is to reduce the incidence of hazardous actions and traffic violations. The automated speed measurement and human observations are designed to produce a direct measurement of changes in the undesired behavior. Consequently these measurements constitute the fundamental information for evaluation.

Driver Survey

An at-home survey of licensed drivers is recommended to complement the other data gathering activities in the Increased Enforcement Activity program. It was suggested earlier as a means of assessing the immediate effect of the publicity campaign, but was dismissed at that point as being too expensive for such a limited purpose.

However as a means of evaluating the effect of the entire project, the at-home survey is much more attractive. It can provide a far better understanding of changes that occur during the project. The survey can help to indicate whether the violation patterns observed over the course of the project were indeed associated with the activities of the project or were due to some other cause. The questions the survey seeks to answer are: "Did the activities of the program change drivers' awareness of law enforcement?", and "Did this changed awareness change driving habits?".

The survey can also help to measure other pertinent questions which concern the overall evaluation. Possibly, substantial changes in driving patterns can be detected· driving pattern in this context refers to where, how much, and why driving was performed rather than the skill with which it was done. These changes in driving patterns will help the evaluator understand what changes have occurred. Furthermore, the survey can measure public acceptance of the program which can be a key factor to local officials.

Since the home interview technique is quite expensive, it may be desirable to combine the interviews with those of another agency which desires unrelated information concerning residents of the city. Additions of extra questions to existing surveys do not substantially increase their costs, but sharing of interview expenses can make the joint effort worthwhile to two agencies where the full cost cannot be justified by either alone. Such possibilities must be explored.

Like all measurement tools, the survey is subject to inaccuracies and will not necessarily be able to do exactly what is desired. Given the peculiarities of the human psyche, it is not possible to ask people directly and in detail what types of changes they have made in their behavior as the result of a specific program. People often may be unaware of their behavior especially when it involves habitual acts such as driving. Further, people are likely to report their actions as being socially desired behavior; this is particularly true when other forms of behavior may be illegal. In conducting such surveys, indirect approaches must often be used, thus producing inexact measurements. Further, such surveys are always subject to question as to how well the subjects represent the population as a whole.

It is proposed that a survey of 500 persons be taken immediately before the project begins, six months after it begins, and one year to eighteen months after it begins. Under this plan, drivers will be selected at random from license lists for the area. The data from the responses will be processed and analyzed to see if critical differences in awareness and reported behavior exist over time. The sample size will be sufficient to detect large shifts in the population as a whole and in certain broadly defined sub-groups of the population.

A suggested list of questions is presented in Figure 2-3. These items have not been pre-tested, but a pre-test will be necessary before the survey is conducted. It may be desired to retain an organization which is experienced in the operation of surveys.

The driver survey, if properly conducted, can provide the final step necessary for a complete evaluation. As has been discussed, a set of immediate goals can be measured for each activity, i.e. number of arrests for training. Further, changes in violation experience can be detected with high accuracy. Still, given that each activity meets its immediate goals, one must make the assumption that achieving these immediate goals is causally related to the change in violations. This assumption is plausible, but its validity must be carefully examined considering the many other unmeasured factors which could have produced the observed effect. The driver survey provides the examination of the assumption by trying to link the immediate goals with driver reactions to violations.

2.3.4 The Defensive Driving Course

Chattanooga has proposed to use the National Safety Council's Defensive Driving Course (DDC) as a court treatment for drivers convicted of certain traffic violations. The long term objective, of course, is to raise the average level of performance of drivers in the Chattanooga area and consequently to reduce the number of accidents and injuries. More

FIGURE 2-3. POSSIBLE DRIVER SURVEY QUESTIONNAIRE

This questionnaire is being conducted as a part of a study of city services in (city name). Your answers will be very useful to us in planning for future city improvements and other services. All answers will be held confidential.

(1) Your name:____________ (2) Your sex:___________

(3) Your address:____

(4) Your age:___________ (5) Your occupation:_________

(6) Length of time at this address:_____yrs._____mos.

(7) Length of time live in (city name):______yrs.______mos.

(8) Do you work in (city name): Yes No

(9) Do you regularly drive a car? Yes No

(10) How many miles did you drive last year?__________

(11) What is your best estimate of the miles that you drove in (city name)?____________

(12) How many miles do you drive every week to and from work and in the cause of business?_____________

(13) How many miles do you regularly drive every week in shopping, going to or from schools or churches, and other personal business such as doctors appointments?_____________

(14) How many miles do you drive every week on recreational journeys such as to parties, to shows, and to parks?________

(15) Shown below are a list of major streets in (city name). Would you please indicate how often you use these streets by checking under the appropriate column.

(List of STEP streets and of control streets)

(16)

Street Name:	Never	Very rarely	At least once a Month	At least once a week	At least once a day

(For each of those streets which the respondent stated that he used once a week or more often he should be asked to rate the following items)

(17) Do you feel that the speed limit on (street name) is:

Much too low Too low About right Too high Much too high

On (street name) do you notice that the number of traffic violations compared to your usual driving experience are:

Much more than average | Somewhat more than average | About average | A little less than average | Much less than average

(18) On (street name) do you find traffic regulations that are more of a nusiance than a benefit to motorists.

Yes, quite a few. No.
Yes, a few. No and they could use some more.
Yes, but only one.

(19) On (street name) do you usually:

Pass most cars.
Pass more cars than pass me.
Pass about as many cars as pass me.
More cars pass you than you pass.
Most cars pass you.

(20) On (street name) have you noticed that

There are many more police patrols than in other parts of the city.
There are somewhat more police patrols than in other parts of the city.
There are about the usual number of police patrols than in other parts of the city.
There are somewhat fewer police patrols than in other parts of the city.
There are a many fewer police patrols than in other parts of the city.

FIGURE 2-3 cont'd.

(21) How often in the past week have you seen the following traffic violations committed? Indicate the best estimate that you can make.

Speeding_____
Ran stop light_____
Ran stop sign______
Cut in and out of traffic______
Jumped light or burned rubber______

(22) In the past year it seems that

Many more of my friends have been given traffic tickets than in the past.
A few more of my friends have been given traffic tickets than in the past.
About the same number of my friends have been given traffic tickets than in the past.
A few less of my friends have been given traffic tickets than in the past.
Many fewer of my friends have been given traffic tickets than in the past.

(23) Have you been taking longer to get to work in the morning in the past few months than before.

Yes, much longer.
Yes, a little longer.
About the same.
No, a little shorter.
No, much shorter.

(24) In the recent past have you noticed that your friends are:

Much more careful in obeying traffic laws.
Somewhat more careful in obeying traffic laws.
About the same as always in obeying traffic laws.
Somewhat less careful in obeying traffic laws.
Much less careful in obeying traffic laws.

From time to time almost all drivers will slip up and not fully observe the traffic law. Before we have asked you to consider your friends and others actions. We now ask that you carefully consider your own behavior and respond accordingly.

(25) In the past week, was there any time in which you felt that your behavior was not in keeping with the traffic laws?

Yes No

(26) In the past week would you say that you drove faster than the posted speed limit.

Not at all. Three times.
One time. Four or more times.
Two times.

(27) In the past week would you say that you went through on a yellow light.

Not at all. Three times.
One time. Four or more times.
Two times.

(28) In the past week would you say that you did not stop completely at a stop sign.

Not at all. Three times.
One time. Four or more times.
Two times.

(29) In the past week, would you say that you made an improper turn.

Not at all. Three times.
One time. Four or more times.
Two times.

(30) In the past week, would you say that you went the wrong way on a one way street.

Not at all. Three times.
One time. Four or more times.
Two times.

FIGURE 2-3 cont'd.

(31) In the past year have you been stopped for a traffic violation?
No. — Two times.
One time. — Three or more times.

(32) If you were stopped at all in the past year, when was the last time you were stopped.
Today. — Sometime in the last three months.
This week. — Sometime in the last six months.
This month. — Sometime between six months ago and one year ago.

(33) If you were stopped at all in the past year, what action did the officer take when he stopped you.
Gave me a verbal warning.
Gave me a written warning.
Gave me a ticket.
Took me into custody and transported me to jail.

(34) If on this last occurence you were issued a ticket or taken into custody which of the following happened.
The case is still pending.
I forfeited bond.
I paid a fine without appearing in court.
I appeared in court, pleaded guilty and was fined.
I appeared in court, pleaded guilty and was both fined and sent to jail.
I appeared in court, pleaded guilty and was both fined and had my license suspended.
I appeared in court, pleaded not guilty and was acquitted.
I appeared in court, pleaded not guilty and was convicted and fined, jailed, or had my license suspended.

(35) On this last occasion, with what were you charged.
Speeding. — Careless driving.
Running red light. — Reckless driving.
Running stop sign. — Driving while intoxicated.
Improper turn. — Defective equipment.
Drove wrong way on one way street. — Other offense.
Following too closely. — More than one of the above.

(36) On the last occasion, would you say that the officer who stopped you was
Very courteous and professional.
Correct in his conduct but not friendly.
Neutral or indifferent in his conduct.
Somewhat impolite about the whole thing.
Very discourteous and unprofessional.

(37) On the last occasion, would you say that the officer who stopped you was
Correct in stopping you.
Fair and honest in his decision.
Used his judgment which might have been an honest difference in opinion.
Somehow or other trapped me in a momentary lapse.
Arbitrarily picked on me without any cause.

(38) On this last occasion did the contact come about through an accident? Yes No

(39) Do you now feel that if you had known more about the law or had a good lawyer you could have avoided punishment? Yes No

(40) What does the phrase "(use program slogan)" mean to you. (Interviewer will characterize answers on the following scale: 0-Don't know
1-Incorrect response
2-Some city program
3-Some police department action
4-Selective Traffic Enforcement Program)

FIGURE 2-3 cont'd

(41) Where did you first here this phrase
(Interviewer will characterize answer as the following.)

1-Just now	5-Signs or posters
2-Newspapers	6-Bumpter sticker
3-Television	7-Friends
4-Radio	8-No response

(42) Do you believe that, in the past year, the police in (city name) have been·
Much more strict in enforcing traffic laws.
Somewhat more strict in enforceing traffic laws.
About the same as in the past in enforcing traffic laws.
Somewhat less strict in enforcing traffic laws.
Much less strict in enforcing traffic laws.

In forming this judgment did you rely mostly on

Observations of police actions taken against others
Reports of my friends and associates.
Experience of my family members.
My own experience.

(43) While driving in the past year have you been
Much more worried about being stopped than in the past.
Somewhat more worried about being stopped.
About as concerned as ever.
Somewhat less worried about being stopped.
Much less worried about being stopped.

(44) Has your present level of concern come about because
You changed your driving habits.
The police have changed their enforcement.
Both change in habits and in the police.
Neither change in habits or in the police.

specifically, the purpose is to improve the performance level of those drivers who have been exposed to the course, since they have been identified as problem drivers and are presumably more likely to be involved in an accident.

Evaluation of this activity represents a distinctly different problem than that of measures aimed at the general population. General measurements are unlikely to detect changes attributable to this activity since the target group represents a small fraction of the total population. While the group toward which the Defensive Driving Course is directed accounts for a large number of violations, this number can still be a small part of total violations. Consequently, evaluation of the course has not been included in the previous discussion, and this evaluation can be considered independently.

The state of the art of evaluation of driver education and driver improvement programs has been the subject of numerous conferences and discussions, and the consensus has been that there have been few adequately controlled studies. Recent work by Coppin et al in California has been good, and there have been several other programs in which appropriate controls were established and used. With respect to the Defensive Driving Course, we know of no such controlled tests. And as a result, we cannot say with any certainty that the course has value as a driver improvement treatment, nor can we say that it does not have any value.

To define success for the DDC requires information in four areas: (1) the normal behavior of the treated group, (2) the type of the control group, (3) the risks that the decision maker is willing to accept, and (4) the minimum average level of improvement considered necessary for the program to be cost-effective. Complete elaboration of the statistical problems involved in these areas is presented elsewhere in this report and will not be repeated here; each of the four areas is discussed briefly.

(1) Suppose that the group selected to take the defensive driving course, displayed very poor driving performance for a year prior to the course. Two conditions can exist: (a) these individuals are habitually poor drivers, or (b) these individuals are average drivers who happened by chance to have a bad year. Under the first alternative, even a very small improvement in the individual's accident or violation experience can be considered a success. Under the second alternative, the course can be only considered a success if the group has an accident or violation rate which is even lower than the population average; their statistical expectancy is to have an average experience in later years.

(2) Three types of control groups can be used; a randomly selected control which receives no unusual treatment, a randomly selected control which receives a different treatment, and a before and after comparison of the DDC group's own record. Use of the untreated control group is most desirable since the results of the evaluation ought to indicate clearly the absolute performance change attributable to the DDC. Given that an unambiguous measure is available, even small improvements in performance can be considered a success. Application of two different treatments to two groups will only tell which program does relatively better. It will not yield information on whether either or both treatments were better than doing nothing. Use of the before and after technique is subject to the problems of regression to the mean alluded to in Reference 2-1. The base period must be sufficiently remote in time so that any short-term chance effects will not be present, but the base cannot be too remote, lest such underlying factors as age and driving experience will have changed. Further measurement of violations in the post-treatment period can be contaminated by the other efforts of the program to increase apprehension of violators. Considering these

problems, much more stringent success criteria should be applied to before and after measures than are applied to a trial with an untreated control group.

(3) The decision maker must choose what mix of risks he is willing to assume. He is confronted with the possibility either of halting a program which might be cost-effective of or continuing a program which is not cost-effective. At one extreme, he can demand a very stringent success criteria but with a high risk that he may disapprove a valid effort. At the other extreme, he can accept a weak indication of success in hope of accomplishing good, but at a high risk of committing scarce funds to wasted effort.

(4) Finally, the success criteria must be established on the basis of what the decision maker considers to be an acceptable increase in performance in cost-effectiveness terms. Whether a 5%, 10%, 15%, or greater reduction in accidents or violations is considered a success, depends on the assessment of the program cost and of the value to the community of the particular reduction.

It is anticipated that the principal measures of success for the course will be based on the students' subsequent violation and accident experience. Changes in the accident experience represent the long-term criteria discussed earlier, and changes in the violations are the intermediate criteria. An immediate measure can be the driver performance measures suggested in Section 4 of this volume (Road User Preparation).

The Chattanooga program offers an opportunity for a controlled test of the Defensive Driving Course, but it will require some rather dedicated cooperation of all parties concerned with the program to insure the statistical validity of the results. If such cooperation can be assured, it will be worthwhile to proceed with a careful evaluation; if not, then it will not be useful to do more than count the students participating in the course just as a measure of activity.

Several years ago a program for treating convicted drinking drivers was to be experimentally evaluated in a large western city. It was agreed that those convicted would alternately be assigned either to a (1) group therapy program or (2) jail term. Assignment was to be essentially random, but the traffic court judge reserved the right in extreme cases to choose which method of treatment should be used. This was expected to happen very infrequently. As it turned out, however, every case was decided on its merits, and the judge could not bring himself to assign randomly. This may have been quite appropriate from an operational point of view although the experimental designer concluded that an evaluation could not be made. The point is that if one is going to obtain a statistically acceptable result, biases of this sort must be excluded from the experiment.

In the case of the Defensive Driving Course, it should be possible to find an alternative treatment which would be acceptable to the courts, a different course, a larger fine, or an assigned probationary period (without a formal course). Assignment to each group can be based on some characteristic not under the control of the arresting officer or the judge. The last digit of the traffic ticket serial number (odd or even) would be appropriate. The use of an untreated (except for the normal fine) control group would be preferred from an evaluation viewpoint.

Chattanooga has indicated that about 600 violators per year will be included in the DDC program. The criteria for inclusion is that a person must have committed two violations in a year. In a recent unpublished study of a similar course in Michigan, participants averaged 0.7 violations a year except for the year in which they took the course when they averaged 1.7 violations. If one assumes that the 0.7 violation rate will hold for persons eligible for the course in Chattanooga, then one will expect the untreated control group to have 420 violations in the year following the course. If the students have 401 or fewer violations then one can be reasonably certain, i.e. nineteen chances out of 20, that the course has improved the students. If, however, the control group displays the usual 0.11 violations per year of the normal driving population, then the students must have 53 or fewer violations in the following year to have the same probability that the improvement did not occur by chance. Since accidents are somewhat less infrequent than violations, a proportionately greater reduction in accident experience must occur for the results to be statistically significant. The point is that the degree of improvement necessary to assure statistically significant results depends on the regular behavior of the students before being treated. Without an adequate control group it will be impossible to tell if there has been an improvement. Table 2-2 presents the estimates for both accident and violation experience under the assumptions that the Chattanooga students come from either a normal or a poor driving group. The estimates are based on the normal approximation of the binomial distribution with N=600.

TABLE 2-2. ASSUMED ACCIDENT AND VIOLATION EXPERIENCE DDC STUDENTS VS CONTROL GROUP

Parent Popul.[2]	Control Accidents	Student Accidents[2]	Change[2]	Control Violat.	Student Violat.[2]	Change[2]
Normal drivers	36	26	10	66	53	13
Problem drivers	285	265	20	420	401	19

[1]Parent population is limited to the type of driver that controls and students resemble in accident and violation rates.

[2]Numbers represent the approximate frequencies necessary for the students' experience to be significantly less than the controls at the 95% confidence level.

Now suppose that the decision maker desires more from the evaluation than statistically significant results. There may be some amount of accident or violation reduction that he wishes to achieve to make the course cost-effective. On the average, these reductions will occur if the course works as desired. In any one year though, reductions of different sizes than the desired size can occur and probably will. The distribution of these reductions will depend both on the desired reduction and on the parent population. For the evaluation, the decision maker must set a level of performance to be achieved if he is to continue the course. In selecting this level, he must balance two types of risk: (1) the observed reduction can fall below this level even though the program will on the average meet his success criteria; (2) the observed reduction can be greater than the minimum level even though on the average the program does not meet the success criteria. For example if he desires a reduction of 45 violations per year from the poor driving group, accepting the program if as few as 17 fewer violations is observed will reduce the chance of discarding a cost effective program to 5%, but such a decision will produce approximately a 95% chance of accepting a program which is not cost effective. Conversely, requiring the program to produce a reduction of 73 or more will yield a 5% probability of accepting a program which is not successful, but will produce approximately a 95% probability of rejecting a successful effort. Table 2-3 presents the data on these ranges for accidents and violations of both the normal population and the poor driver population.

In summary, there are a number of statistical problems associated with evaluating the Defensive Driving Course. If students in Chattanooga are similar to those in Michigan, we will be able to determine with some confidence if the course yields a 7% reduction in accidents or a 4.2% reduction in violations. A highly controlled test of the value of the Defensive Driving Course seems in order in the Chattanooga program, but if proper controls cannot be established and maintained there will be little purpose in an evaluation.

TABLE 2-3. NECESSARY ACCIDENT AND VIOLATION REDUCTIONS FOR BOTH NORMAL AND POOR DRIVER POPULATIONS IN DDC.

Group	Crashes:[3] Low Level[1]	Crashes:[3] High Level[2]	Violations:[4] Low Level	Violations:[4] High Level
Normal drivers	19	31	35	55
Poor drivers	13[(5)]	53	17[(5)]	73

[1]Low level represents the amount of reduction necessary for a 5% probability that the goal has been achieved.

[2]High level represents the amount of reduction necessary for a 95% probability that goal has been achieved.

[3]Goal is a reduction of 25 accidents.

[4]Goal is a reduction of 45 violations.

[5]Not a statistically significant reduction.

2.4 A National Laboratory Plan

Several benefits will accrue from the several federally funded change programs in the law enforcement field, particularly if useful measures of effectiveness are available. The results will of course be of interest to the jurisdiction conducting the operation, as the local government will have to decide whether to continue the work after federal funds have disappeared. How well the results can be extrapolated to other communities is not evident; the change may have been specific to the site from either a geographical or social point of view. Nevertheless, it can be expected that observed results will be of some national value.

We suggest, however, that there be instituted two programs which will constitute a national laboratory for the conduct of and evaluation of experimental changes in the law enforcement programs. Such laboratories will provide structured tests of many programs under similar conditions, a feature that is not readily available from scattered site demonstration projects.

There have been in the past such national experimental sites for other reasons--several highly instrumented freeways (in Chicago, Houston, and Detroit) were considered in that light, and in the public health field the town of Tecumseh, Michigan is engaged in a long-term program of epidemiological measurement of health conditions over time in a highly observed population. The NHTSA Alcohol Safety Action Program also serves as a prototype for such experimental programs in a community setting.

As indicated earlier, most communities institute changes in this area in a slow and uneven manner. More often than not these changes are not fully planned nor are they carefully evaluated. This state of affairs persists not because the communities are either inherently unable or unwilling to do a better job. Rather with the multitude of problems confronting them, they lack the time and the resources to explore fully any problem area. The intent of the national laboratory plan is to attack these problems through a set of carefully planned, executed, and evaluated programs. This activity will be of value in two ways: (1) after the evaluation of several programs, the laboratory can serve as a guide to many communities in choosing effective programs to solve their problems, and (2) the process of well conceived action of the laboratory can demonstrate to other communities the clear value and long-term savings obtained from planned and systematic change.

2.4.1 Sites

In selecting the sites for the laboratories, three criteria must be applied. First, the community will be representative in some way of many communities experiencing highway safety problems in the law enforcement and other areas. Second, the community must be receptive to being used as a test area; previous success in conducting experimental programs will weigh heavily in an area's favor. Third, there must be available in the local area sufficient talent both to operate and to evaluate the efforts.

Cities of the order of 100,000-200,000 residents have the fastest growing traffic fatality problem in the nation. Over the past decase, from 1960 to 1970, traffic fatalities have risen everywhere, but nowhere at a faster rate than in these cities. One can only speculate as to the reasons. These have generally been the expanding cities, taking in more suburban territory; evidently they have the most difficult traffic problems of both the small town and the major urban area.

The city of Flint, Michigan, has had an active program in police traffic services sponsored by NHTSA funds over the past several years. It serves well to represent the urban area described here, and has certainly shown a demonstrated interest in attacking the traffic accident problem. The present program there, however, lacks detailed instrumentation; and as a result it is difficult to trace the effects of particular countermeasures.

Monroe County, Indiana, has had an active NHTSA-sponsored police experimentation program for the past two years. Monroe County is perhaps a typical midwestern agricultural community, with a county seat (Bloomington) approximately at the geographic center. In this particular case, however, Bloomington happens to be the home of the Indiana University Institute for Research on Public Safety which has put a considerable effort into instrumentation of the county specifically for the purpose of evaluating changes in the law enforcement practices of the community.

Both Flint and Monroe County, then, have a demonstrated interest in experimentation in the law enforcement area; Flint has an active police department with a proven history of performance in traffic services; and Monroe County has an ongoing instrumentation system as well as three police agencies (state, county, and city) who have participated in experimental programs in the past.

We propose, then, that these two communities be considered for selection as continuing, long term, national demonstration area in which there can be a sequential implementation of changes in the law enforcement program. These changes will be of national interest and widely applicable, but of local value to the particular communities at the same time.

Both communities will be brought to the same degree of instrumentation so that there will be a permanent capability to measure and record information about both traffic flow (speeds, headways, violations) and accidents.

2.4.2 Program Phases

There are four phases to the demonstration programs suggested for these two communities and all four must be more or less continuously in operation. They are:

(1) The definition of problems for solution.

(2) The planning of change programs responsive to the problems.

(3) The conduct of the programs themselves.

(4) The measurement and evaluation of the programs.

While these are the same phases which are present in nearly any program, they differ in the present case because of the national significance. This should be particularly apparent in the problem definition phase, and in the nature of the evaluation and reporting.

The strategy to be followed in the laboratory is one of following programs through all four phases of activity in a carefully observed and measured manner to produce results that have general applicability. Two types of programs will be conducted: major system changes and detailed analysis of specific techniques. Major system changes will be those activities which greatly alter the nature or extent of activities in one or more of the eight countermeasure groups, e.g. a doubling of manpower devoted to traffic patrol, or a restructuring of the court system. The concept for major system changes is one of incremental change. At each step, the countermeasure groups will be examined and the area which offers the possibility for the most substantial payoff will be selected for implementation. In implementation, the best techniques in this area will be applied to the specific problems in the communities. The results will then be analyzed. Then the countermeasure group with the next higher payoff will be implemented. The analysis of specific techniques will be of a more limited nature and will seek to determine the best means of conducting a particular activity; a good example is an evaluation of the relative merits of radar vs Vascar vs Orbus in detecting specific types of speeding offenses. The major system change portion will seek to determine the most cost-effective level of activity in a particular area while the analysis of techniques will seek to determine the most efficient means of achieving the desired level of activity.

2.4.3 Problem Definition

Eight areas of possible action have been identified in the area of road user regulation. Among these eight areas, there exist a large variety of possible programs to be implemented. Therefore, critical task in the laboratory concept is to pick a particular set of programs to test in the communities. The criteria for selection will include the nature of problems present in the areas, the degree of innovation in possible programs, the national relevance of the results, and the degree to which the programs serve to round-out other research.

Guidance in selecting the projects for implementation will come from two types of advisory panels. The first type of panel will be nationally constituted and will oversee both community activities. The second type of panel will exist in each of the communities and will serve to relate the suggested programs of national significance to local conditions.

The national advisory committee will include prominent individuals from the law enforcement, legal, court, scientific, and highway safety communities. The chairman should be a law enforcement official of national reputation. Program directors from each community will serve as ex-officio members of the group. The advisory committee will review proposed programs for their general applicability and suggest those changes (countermeasures) most usefully carried out in the instrumented environments of either location.

The local advisory committee will be constituted from approximately the same fields as the national committee, but the members will generally be operating heads of the responsible local agencies. Citizen participation will also be most helpful at this level. The primary role of the local committee will be to identify local problems to which the suggested national programs can be applied and to coordinate the efforts of the operating agencies in executing the program. The chairman of the local committee will be the head of some local organization which will be heavily involved in actual operations.

The national committee will meet at approximately six month intervals, and the local committee chairmen will be invited to participate in these meetings. The local committees will meet monthly, and their purpose in these meetings will be to insure the continued smooth operation of the programs.

2.4.4 Planning

Once particular activities have been selected for implementation, a detailed operational plan and a complete experimental design must be created. This activity will be the responsibility of a professional planning staff which will also have responsibility for conducting the evaluation. This staff will consist of a core of personnel drawn from the operating agencies and of scientifically trained persons for experimental design and statistical evaluation. The personnel from the operational agencies will be responsible for integrating the experimental activities into their agencies' ongoing efforts. The scientific personnel will be responsible for the design and conduct of evaluation measurements. From time to time outside experts in both the operational and scientific areas will be included for specific advice and consultation.

The philosophy of evaluation of demonstration programs is covered at length elsewhere in this report. Nevertheless it may be worthwhile to repeat the thought that the purpose of careful measurement in the evaluation process is to permit the experimenters to find out not only that some countermeasure "worked", but to discover why it worked. When faced with a successful program we are bound to ask the question "What did we do right?", and the principal purpose of a continuing national experimentation activity is to insure that the answer to that question will be available.

Each plan will be specific to the particular activity which is being implemented at the time, but will need to be tied to the other plans through a common approach. It will be most useful to the planners to keep in mind the causal chain concept which is discussed in detail elsewhere in this report. The planners should attempt to relate each activity to the three major links in the chain in the regulation area: driver characteristics, driver performance, and system performance. By so connecting the individual plans to an overall program concept, the planning staff will insure both practical results for the immediate effort and a contribution toward solving the overall problems associated with regulatory activity.

The planning staff, then, must take the recommendations of the advisory bodies, and turn them into working plans--detailed as to time, cost, measurement, analysis, etc.

2.4.5 Operations and Implementation

Actual operations will be conducted by the local police, courts, and other agencies in accord with the plan devised by their representatives on the planning group. Until the actual programs are established, not too much can be stated about actual operations, but several cautionary suggestions are offered.

When feasible, the operational and measurement personnel should come from distinct groups, perhaps not even within the same organization. Both types of activity require specialized knowledge and abilities which often require the full-time attention of an individual. This is particularly true of the police officers engaged in the experimental activity; to use them on data gathering tasks will not make efficient use of their law enforcement abilities. One can imagine the demoralizing effect of asking a trained police officer to stand idly by recording data while hazardous violations are being committed. Also an individual's desire to perform well in an operational role can possibly detract from his objectivity in a measurement activity.

The activities of the operating agencies will need to be closely monitored so that evaluation can be based on actual performance of the organization rather than on assumed activity levels. Program officials must be alert to subtle, and perhaps not so subtle, changes in policy and practice particularly at the lower supervisory levels which can alter the performance of operations.

Many of these problems can be avoided by including some operational personnel on the planning staff. If the plans are drawn by people with experience in the operation, many pitfalls can be easily avoided. Further, extensive effort should be made to enlist the wholehearted cooperation of operating personnel.

2.4.6 Evaluation

Most of the tasks and evaluation activities for the laboratory can be conducted at any site. Concentrating these activities in two locations over an extended period of time produces several desirable effects for evaluation. Long-term results of any particular activity can be observed. The incremental approach to conducting activities can be used with the advantage that short-term variations not under the control of the project can be recognized and their effect on the particular activity observed. Sophisticated data gathering techniques can be used which might not be justified for a single project over a short time. Finally, as experimental and observational activities become common place, the performance of individuals in the system will come to approximate their normal response to the programs rather than a reaction to the observers.

Most of the evaluation techniques have been discussed in detail in the context of the Chattanooga project, and repetition is not necessary here, except to stress the importance of having reasonably unobtrusive measurements which will reflect peoples' true behavior. The differences between the measurements in the laboratory concept and those discussed earlier is one of degree.

The loop detector network will be greatly expanded over the twelve sites recommended for Chattanooga. Further, it is anticipated that continuous monitoring will be undertaken at a number of the sites to give an ongoing record of trends in traffic and speed variations.

The recording of other violations such as illegal turns and improper lane usage will be somewhat expanded. It may be feasible to permanently install television cameras at several key locations so as to monitor performance at these sites on a much more frequent basis than is envisioned in the earlier section. Further, such techniques as time-lapse photography and other detection devices for the non-speed related offenses can be employed and refined.

A continuing program of monitoring non-experimentally controlled conditions will be maintained. Information should be regularly and systematically gathered on items such as driving exposure and experience of the population, local economic conditions, changes in the vehicle population and driving records. Detailed and accurate measurement of the activity levels of all agencies associated with the project must be maintained.

The program of at-home interviews of drivers to test public reaction to the program will be made a permanent feature. By so doing, the immediate reactions to each new program can be determined and the cumulative effects of all the programs can be observed. Special interviews will also be employed to inquire into specific problems which arise such as the impact of a public information campaign or the amount of time necessary for people to become aware of a new traffic regulation.

It is anticipated that the evaluation staff will also devote some portion of their effort to devising better evaluation techniques. Also the staff will be encouraged to devise new and more powerful predictive models of the likely effects of particular actions as their understanding of the data suggests new hypotheses.

The ultimate goal of the evaluation effort will be to determine which of the programs and techniques prove to be the most cost-effective way of reducing the loss from crashes. Given well conceived and adequately measured projects, this goal can be attained.

2.5 Cost Estimates

Costs of the Road-User Regulation demonstration program were estimated on the basis of the previously mentioned STEP proposal for Chattanooga, Tennessee, and existing programs in Flint, Michigan and Monroe County,Indiana. In the latter two locations (the national laboratories), the estimates are given as ranges of costs because of the flexibility in program choice to be exercised by the local committees.

It is assumed that all three areas will have very active programs for at least three years, though the national laboratories should probably continue longer if possible. It is also assumed that thorough traffic flow instrumentation will be installed in all areas, and that computer facilities will be purchased and amortized over five years. Finally, it is assumed that thorough evaluation will be performed; in Chattanooga it will cover only the four specific program areas independently, while in the other locations the evaluation will consider a much wider range of program areas, and their combinations.

Table 2-4 presents the annual cost-component estimates for three year programs in each location. The differences in operational costs for Flint and Monroe County are due to their different sizes. The grand total of costs would range from $2,700,000 to $4,200,000.

TABLE 2-4. ANNUAL COST ESTIMATES FOR ROAD USER REGULATION DEMONSTRATION PROGRAM

	Operational Costs	Instrumentation and Evaluation Costs	Total Costs
Chattanooga	$250,000	$50,000	$300,000
Flint	$150,000 to $350,000	$200,000 to $300,000	$350,000 to $650,000
Monroe County	$ 50,000 to $150,000	$200,000 to $300,000	$250,000 to $450,000

2.6 Summary

The general area of road user regulation has been subdivided into eight countermeasure groups which provide a functional way of looking at problems of implementation and evaluation. Four of these eight groups have been selected for detailed analysis in the context of a particular city, Chattanooga, Tennessee. In each of the four areas a set of objectives-immediate, intermediate, and long-term was suggested and the techniques for measuring the achievement of these goals were outlined in some detail. While the coverage of areas was not complete and other evaluation techniques might be suggested, the countermeasure areas and evaluation techniques examined are those of most common use and general applicability.

We have concluded that any attempt to implement every possible combination of the countermeasure groups presents an unattainable goal. No such experimental plan can be undertaken since it is impossible to obtain enough similar communities so that other differences among these communities will not confound any experimental results.

Rather than attempting such a grandiose and hopelessly confounded plan, we suggest that there be instituted two national laboratory communities in which particular combinations of programs can be implemented on a sequential basis. The particular programs will be selected on the basis of the best judgment available in the field. Once implemented, they will be carefully measured and evaluated to indicate where the greatest payoffs lie. Such a concentrated long-term effort is necessary to insure that the results which are obtained are meaningful and of wide applicability.

Finally, we again stress the need in any program for a program of careful and detailed measurement of activities and particularly of intermediate program objectives. These measurements must be conducted with a fair amount of sophistication before, during, and after operations are conducted, and every attempt must be made to obtain adequate controls. Without such measurements we will never have sufficient information to bring the full weight of the legal system effectively to bear on reducing the toll of accidents.

REFERENCES

2-1. E.F. Fennessy, Jr. and H.C. Joksch. Police Traffic Services and Road Safety: An Evaluation of the Literature. Hartford, Connecuticut: The Travelers Research Center, Inc., n.d. (Prepared under Contract FH-11-6604 with the U.S. Department of Transportation, National Highway Safety Bureau).

2-2. Chattanooga, Tennessee, City of. Selective Traffic Enforcement Program: Proposal. (Unpublished grant proposal by the city.): May 14, 1971.

2-3. A Study of the Effects of Law Enforcement on Traffic Flow Behavior. Bloomington, Indiana: Indiana University, Institute for Research in Public Safety, January 1971. ("Prepared for U.S. Department of Transportation, National Highway Traffic Safety Administration").

2-4. Development of Violation Sampling and Recording Techniques. Washington, D.C: Kelly Scientific Corporation, August 1970. ("Contract No. FH-11-7294, prepared for U.S. Department of Transportation, National Highway Traffic Safety Administration".)

3.0 INFORMATION FLOW DEMONSTRATION PROGRAM

3.1 Introduction

This demonstration program is concerned with the reporting, accumulation and utilization of traffic safety data; i.e. driver, vehicle, highway and accident data. The information flow program is the first step in the long term loop that links "events" of the traffic system to the corrective actions. Thus changes in the information flow system are made for the purpose of effecting improvements in the decisionmaker's ability to modify the traffic system toward greater safety.

While the ultimate criteria for evaluating countermeasure programs is improved safety in the operating traffic system, an information flow demonstration program operates in a somewhat different domain. It might be considered as a separate, self-contained system that directly affects the decisions made in all the other countermeasure programs. The information flow program is analagous to a film studio in which the whole traffic system operates.

The director (who decides how the plot is to be carried out) and stage hands (who take action to change the traffic system) are also onstage. Events are filmed (recorded), accumulated (spliced), analyzed (edited) and disseminated (projected). Normally the decision maker (director) is in the position of a film critic who wants to improve plots in the future. To the degree a decision maker is in this position he is necessarily dependent on the effectiveness of the information flow program.

The purpose of this experimental program is to demonstrate that changes in the recording, accumulation and utilization of traffic records will improve the effectiveness of the data flow. This will be done in two ways:

1. Maximize the utilization of traffic records systems by the local user or highway safety practitioner.
2. Develop exposure measures as a basis for computing accident rates.

The criteria for evaluating system effectiveness is clearly stated in Chapter V of Traffic Records, Highway Safety Program Manual, Volume 10:

> "The effectiveness of the traffic records program is its ability to produce the information needed to support decisions for effective management of the total traffic safety program".

The detailed experimental design which follows contain:

1. Taxonomy of Possible Programs
2. Heuristic Derivation of a Demonstration Program
3. Location of Demonstration Program
4. Location of Demonstration Program Plan
5. Resource Requirements

3.2 Dimensions of Information Flow Demonstration Programs

A taxonomy of the different aspects that should be considered in designing information flow programs was developed in terms of four questions.

1. Where Performed? (Information Flow Sub-Systems)
   Organization - Police, Motor Vehicle, Highway Depts.
   Files - Driver, Vehicle, Highway, Accident Records
   Operational Units - Reporting, Accumulating, Utilization
   Jurisdiction - State, Local
2. How Performed? (System Activities)
   Role - Plan, Control, Operate, Use
   Activities - Training, Programming, etc.
3. How Much? (Intensity/Complexity)
   Level - Plush/Paltry
   Mode - Push/Pull
   Frequency - Day, Week, Month, etc.
   Channels - Mail, On-line Computer Terminals
   Products - Tabulations, Maps, Resource Allocations
4. Who For? (Users/Decision Makers)
   Jurisdiction - Precinct, City, County, State, Federal
   Organization - Police, Highway, Court, Etc.

The first question-"Where Performed?"- addresses the problem of which administrative units implement countermeasures. This facet is concerned with where the money is spent rather than who receives the ultimate benefits of the program activities.

The second question - "How Performed?" - considers the nature of activities performed from planning to computer programming. The first two questions complement each other. Together they define the countermeasure program setting, independent of outcome; i.e., Where and how program monies are expended.

The third question-"How Much?" - is concerned with the intensity, complexity, or level of program. We need to define the output of the countermeasure activities, such as computer produced pin maps. This is another dimension of the first two factors: the amount of money spent.

Finally, the scope is determined by asking "Who For?". The target group for the information flow demonstration is the user or decision maker. We aim at the decision maker's information utilization activities and evaluate the results.

3.3 Heuristic Derivation of a Demonstration Program

A proposed information program can be clearly defined by answering the four questions above. Pairs of these dimensions can be compared in a matrix table, such as "administrative unit" vs. "activities" to see where and how a program would be carried out. The taxonomy, then, is useful in helping to define new programs, to determine where they fit with other programs, and to uncover potential weak areas. We have used the above questions in the following sections as a guide in heuristically deriving the specific countermeasure areas for the proposed information flow demonstration.

3.3.1 Where Performed? (Information Flow Sub-Systems)

Organization

Each countermeasure demonstration program will directly affect specific organizations, government agencies, and departments. In this section of the program we are concerned with those organizations that control and perform information flow functions, not with the users of information. This distinction can be hazy because the source, processing, and use of traffic-record information may be all within the same organization. As stated earlier we wish to improve the flow so as to aid the decision maker or user. To do this, we must spend money to improve the control and performance of information flow functions. But first we must determine where these functions are performed.

There are a number of organizations such as the police, driver license, motor vehicle vehicle registration, and highway departments, involved in processing traffic records. Each agency relies on individual traffic record files and operations; thus, by examining the files and operations we can determine which organizations should be most directly involved in this demonstration program.

Traffic Record Files

The fact that the information flow program is common to most of the other countermeasure program areas can be seen below in the listed types of traffic records.

Traffic Records

- Driver Records
  - Education
  - Licensing/Testing
  - Violations
  - Sanctions
  - Financial Responsibility
- Vehicle Records
  - Registration
  - Title
  - Inspection
  - Stolen
- Highway Records
  - Fixtures Inventory
- Accident Records
  - Police Reports
  - Operator's Reports
  - Bi-Level Studies
- Exposure
  - Traffic Courts
  - Driving Mileage

The primary emphasis of this demonstration program will center on the use of accident records. Through they have the most direct bearing on highway safety, accident record files are probably the least used of all traffic records. Most of the other traffic record systems were established to meet some administrative or legal requirements (e.g. license plate revenue collection) and they generally serve their intended functions well. Thus the highest benefit will be derived in the processing of accident records.

Other traffic records, such as traffic counts, violations, and driving exposure data, will also play a role in the program. The only new type of data proposed involves the reporting and accumulation of exposure data. There are a number of other information sources that could be tapped, such as a driver survey of the worst intersections in town, but these are being considered as part of the other individual countermeasure program evaluations.

Operational Units

Three basic operations can be identified in the sequence of data flow: recording, accumulation, and utilization (which includes both analysis and dissemination.) These operations are usually administered so independently that they can be considered as separate operational units. Each of the traffic record files noted earlier involve these units. Information countermeasure programs can be executed in any one of these operations or in combination, as in Figure 3-1.

Report	Accumulate	Utilize
Report + Accumulate	Accumulate + Utilize	Utilize + Report
Report + Accumulate + Report		

Figure 3-1 Data Operations and Combinations

Reporting Operation

Reporting is basically the operation of recording an event, such as a patrolman filling in an accident report. There is a perpetual problem of accident data inaccuracies in the reporting operation. A major error in existing accident data is due to under-reporting, particularly in the mass of minor accidents.* Another problem is non-uniformity of reporting (e.g., variations in police injury code interpretation among the accidents that are recorded). Some of these biases can be corrected by, for instance, extrapolations of reported totals, using ratios of non-reporting derived from surveys. There has also been a move towards uniform report forms and clearer definition of reporting criteria. The necessity for improved reporting is clear and does not require further demonstration on an experimental setting.

Accumulation Operation

Over $37 million in federal matching funds have been expended in the Traffic Records standard areas. The greatest proportion of these monies were for the planning, development and implementation of computer based data files. The sum is over $50 million dollars when data processing activities in other standard areas are included (primarily in identification and surveillance of accident locations, motor vehicle registration, and police traffic services). A number of states have developed highly sophisticated hardware and software computer facilities. An extensive amount of data has been accumulated in these files, and the operation is well developed. We see no need for further demonstration in this area.

Utilization Operation

In contrast with the extensive efforts to improve reporting and data accumulation, there has been a relatively minor effort in improving the utilization of accident data. The total cost of the information handling facilities in automated state files devoted to traffic records is many millions of dollars. The average, annual cost of inputting accident data alone (from police report to computer

---

* Acquisition of Information on Exposure and on Non-Fatal Crashes, Volume II - Accident Data Inaccuracies, R.E.Scott and P.S. Carroll, Highway Safety Research Institute, The University of Michigan, Report No. 03169-II, Contract No. FH-11-7293, NTIS Distribution, 87 pages, May 1971.

storage) is estimated at more than $4,000,000 per state, Yet it is difficult to identify more than a dozen persons in a typical state who are concerned principally with the analysis and dissemination of this information. The time is ripe for a large demonstration effort in utilization. The data files exist and continue to grow. The purpose of this program is to demonstrate how highway safety might benefit from better use of the data.

Reporting and Accumulating Exposure Information

Two types of exposure information will be reported and accumulated: a statewide survey of driving exposure and traffic counts on local streets. It will also be necessary to add new data elements, such as accident location coordinates, to the accumulation of existing accident reports, and to encourage better reporting of these same data elements within selected experimental communities.

Driving exposure information within classes of drivers and vehicles and roads is essential to the determination of valid accident rates, and to our knowledge it has never been collected on a statewide basis. The number of accidents (accident frequencies) can aid in identifying problems but comparing accident rates, from year to year or between jurisdictions is necessarily dependent on adequate exposure information. For example, if a young driver problem is identified by their high accident frequency among young drivers, some local countermeasure program might be implemented. If during the countermeasure fewer young-driver accidents occur, one would like to know if the drop was due to the countermeasure or a change in the young driver population (e.g. fewer young drivers, driving fewer miles). Therefore a statewide driving exposure survey is proposed as a specific information countermeasure project.

Traffic counts can be used as an accident exposure measure for specific sections of roads. Traffic counts are sometimes restricted to determining road maintenance schedules and predicting future traffic loads (e.g. when to remove parking meters or widen the road). But traffic counts can also indicate density patterns and the mix of vehicles on the road. At an intersection, for example, the exposure information of traffic counts would help to clarify existing accident risks.

Jurisdiction

Finally, we need to determine at what jurisdictional level the work will be performed. While section 3.4 on user coverage will emphasize local decision making, much of the work in providing information services will be at the state level. The large automated traffic record systems are at the state level, because they quite naturally evolved from statewide administrative functions.

The experimental design of the demonstration program calls for three levels of service (in terms of where the work will be performed). Adding jurisdiction to the previous diagram of operational units adds another dimension to the program description, as seen in Figure 3-2. The control level continues to receive existing service. The state level provides added services from a state centered operation and a statewide exposure survey. The local level provides capacity for local reporting, accumulation and analysis of traffic records data. The question of how much service (discussed later) follows the same three levels.

3.3.2 How Performed? (Information System Activities)

The people involved in the Information Flow system play four roles: plan, control, operate, use. The planners and controllers decide what actions are to be taken within the information flow system. The operators perform to carry out the decisions. The users are the recipients of the information flowing from the system and are therefore affected by changes in service. Since one individual may, in time, play all four roles, the distinction among roles is crucial. This program clearly separates the roles of planning, controlling and operation from the user role so that intentional changes in the first three roles are not confused with observed changes in the user's role.

A series of specific change activities (such as computer program preparation) will be executed in order to implement the countermeasure program. Information Flow system changes resulting from these activities are, in fact, the countermeasure itself, i.e. increased data utilization services.

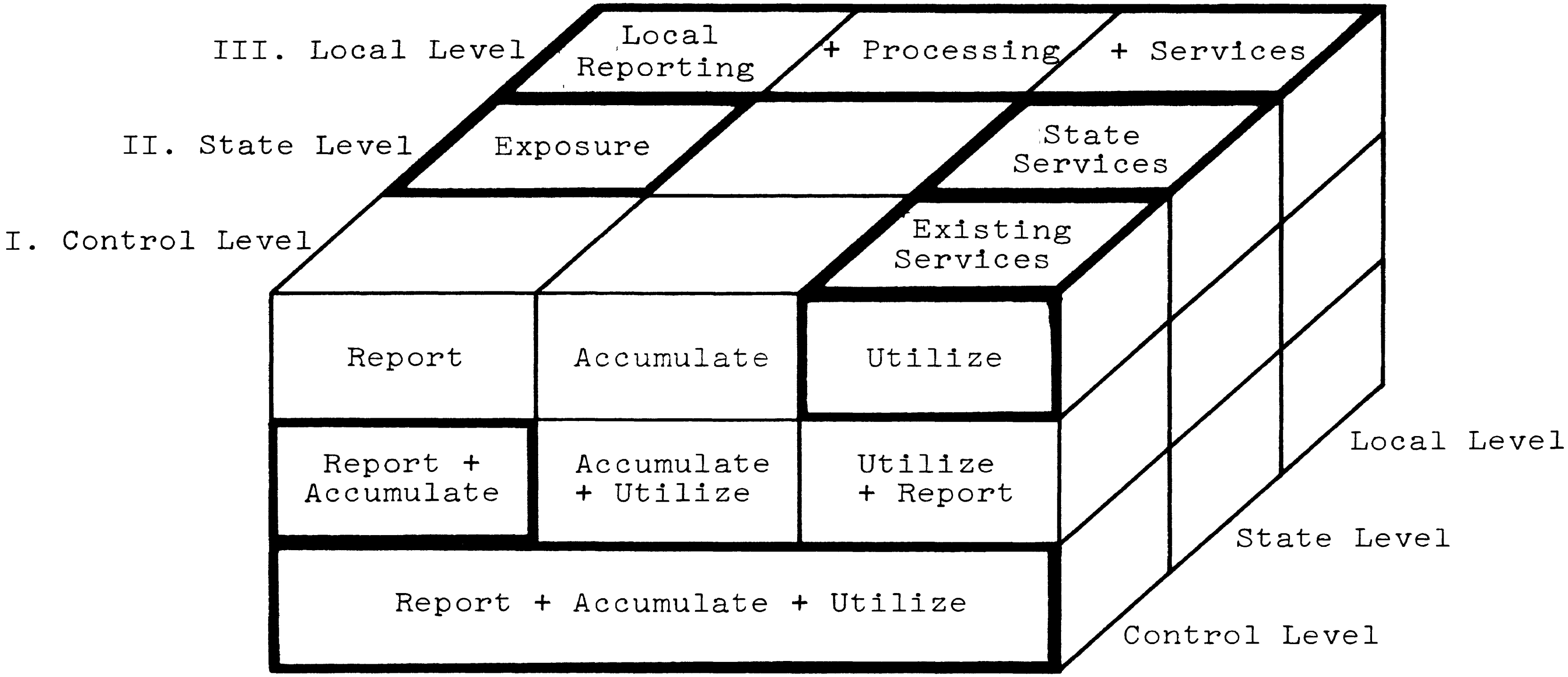


Figure 3-2 Information Flow Countermeasure Program Sub-Systems: Operations and Jurisdictions

Planning and Control

A statewide traffic records committee will be established to improve planning and control of information flow, following the guidelines of the Traffic Records Highway Safety Program Manual, Volume 10. This committee will play a major role in the development of the demonstration program. A user's needs survey will also be conducted to help design the mix of products and services.

Operation and Use

The operation role will be greatly changed, particularly the operation of the information utilization sub-system. Existing analysis and dissemination staffs will receive additional training, new personnel will be added, and computer programs will be prepared. Users themselves will also receive training and indoctrination.

In summary, we will alter the roles of persons involved in the information flow system in order to improve utilization of traffic records, and then we will measure changes in the degree to which the information is used.

3.3.3 How Much? (Intensity, Complexity of Information Program)

Level of Service

We have used the alliteration of paltry, progressive and plush to exemplify the three levels of service. Paltry refers to the existing level wherein few persons benefit from the store of traffic records. In the plush system, we attempt to always get the right information to the right man at the right time. This demonstration program measures the increased benefits of plush levels of service.

Mode of Service

The services provided are a mix of two types: push and pull. A push service automatically distributes products, usually on a periodic basis. The existing monthly summary comes regularly, even if not needed. In a pull service users request the answers to specific questions, usually of immediate concern to themselves. The requests involve the use of individual traffic record retrieval (e.g., driver license status) or aggregate data statistics (wrong way accidents distributed by time of day). Since many traffic record systems are already supported by the demand for finding individual records, the primary emphasis of the program will shift to statistical analysis of mass traffic records and mass accident records. Examples of the types of service are shown in Table 3-1.

Table 3-1
EXAMPLES OF INFORMATION SERVICE

	Type of Service		
	PUSH (Dissemination)	PULL (Demand/Response)	
Level of Service	Accident Tabulations	Driver Record Retrieval	Accident Locations
Paltry	Annual State Accident Summary	Manual Search, Xerox & Mail	List of Accidents at a Location
Plush	Selective Alert of Special/Unusual Community Situations	Court room Video Terminal On-line Retrieval	Pin Map of Selected Local Accidents (e.g. after dark)

Other design considerations such as frequency and communication channels are also factors in implementing the program. For example, a plush level of service may require a more frequent selective dissemination and higher speed data lines for operation of an on-line video terminal. Each of these exemplifies an increase in the complexity and intensity of the program.

The demonstration program will be conducted at three levels as mentioned before. While originally depicted in terms of jurisdictional level at which the work is performed (the control, state and local levels) the program matrix in Section 3.3.1 also describes the levels of complexity or intensity of the program. At the control level, services will be left as they are, i.e., no additional expenditures. The progressive or second level provides for additional services from an analysis staff, centralized at the state level.

This staff will provide many new push services on a selective basis; the right information to the right man. Request or pull services will also be promoted and provided. Staff size and budget will be fixed, and the best mix of services within those constraints will be determined by user needs.

The third level provides the above, plus additional capacity at the local level for reporting, accumulation and analysis of data. For example, accidents could be recorded by patrol and location to aid in police manpower allocation, and additional data elements of value to the local traffic engineer could be recorded. This is the truly plush level of service in which all practical push and pull service modes are available. The relationships between service levels (complexity/intensity) and subsystems (operations) are shown in Figure 3-3.

The full spectrum of services that could be provided are discussed at the end of the information flow system description of Volume 2. The specific products to be provided will be determined by the user. What he wants or needs will determine the best product mix. The first phase of the program involves finding who the users are and what their needs are. An example of a proposed service is given with each user group discussed in the next section.

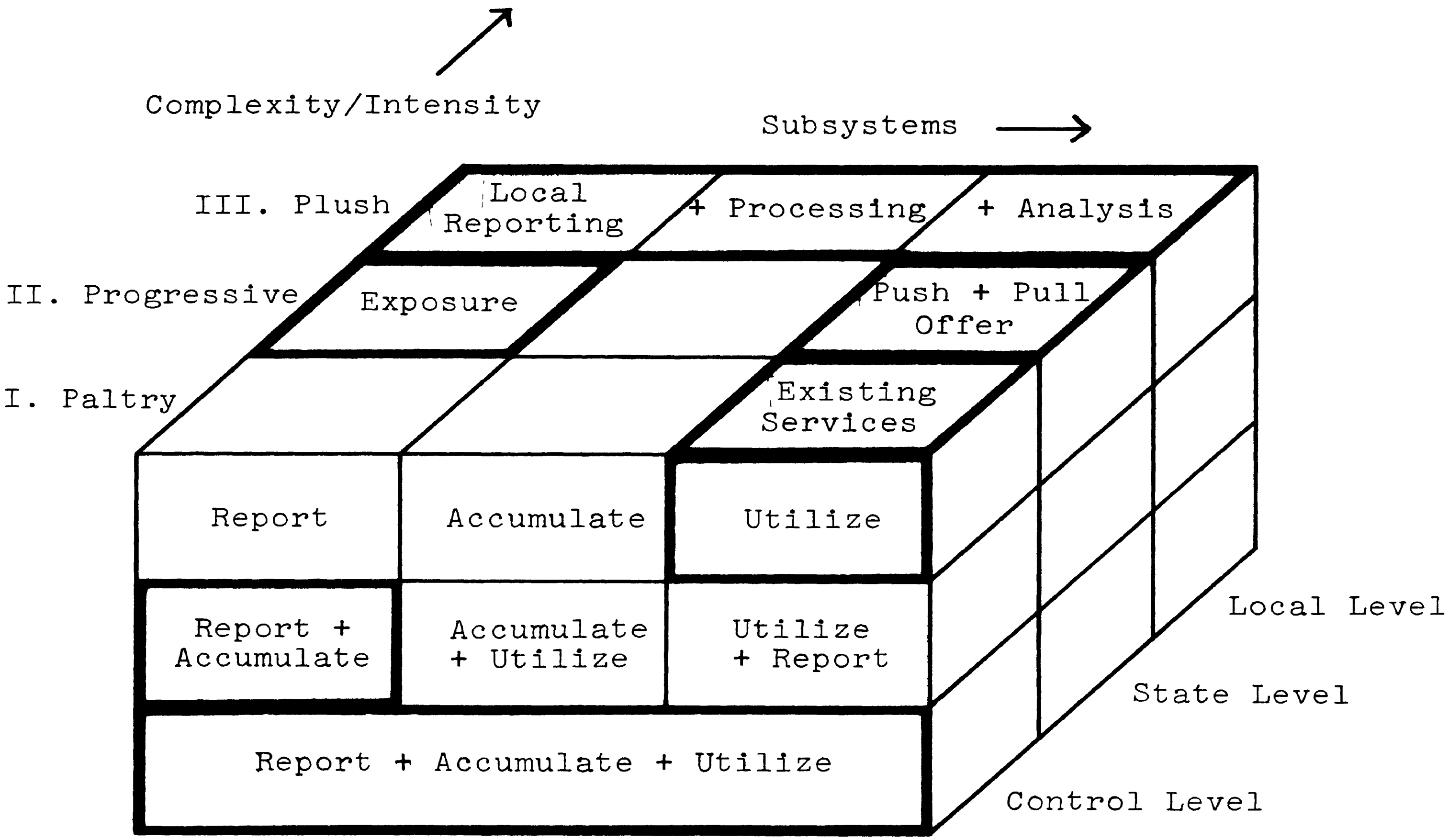


Figure (3-3) Information Flow Countermeasure Program: Sub-systems and Complexity/Intensity.

3.3.4. Who For? (Users/Decision Makers)

While the need for knowing how we are doing makes the implementation of a strong traffic records system imperative, the drive towards implementation can be so strong that large sums of money are spent designing or attempting to design systems in hypothetical situations for undertermined people with unknown needs. Thus, it is very important to clearly identify the users in advance by jurisdiction and function.

Jurisdiction

The target group for the information flow demonstration program is the local decision maker. Past emphasis has been towards state and national accident statistics with little demonstrated concern for the local user. Recognition of the importance of local government participation is clearly stated in the Traffic Records Program Manual, Chapter VII; i.e. local accident problems are best studied at the local level.

In the fully implemented level III, the data flow program requires the establishment of county highway safety teams representing all aspects of the problem. Even with excellent data services, utlization of the services will not increase if traffic safety ranks low on the priority list of local problems. The demonstration program will be successful to the degree that local traffic safety activists (e.g. traffic safety teams and local safety councils) can change the priorities. Information derived from traffic records may help them sell their point.

Organization Affilation

Who are the potential users? The spectrum of people who would utilize traffic records is outlined by the organization affiliation of the users. Each user has unique needs for information and data, so we must also ask; What are his information needs? and Which of these needs should be fulfilled? A user study will be conducted to answer these questions and derive a user/product matrix. The results will also help to relate user's needs to the levels and modes of service of the previous section. Typical user organizations include:

a. Roadway Administrations
b. Police Departments
c. Courts
d. Medical Services
e. Driver Education
f. Safety Councils
g. News Media
h. Legislators

The following paragraphs discuss the user organizations, with examples of information services to be provided.

a. Roadway Administrations

Roadway administrations include county highway departments, city traffic engineers, city transit authorities and public works departments. These organizations make long term, permanent changes to the traffic system. At present, users concerned with state trunklines receive accident location information in order to spot bad road sections. The same ability will be extended to users dealing with other roads in level II and III counties. Level II counties will receive a ranked list or "hit parade" of high accident locations. Level III counties will accumulate X, Y locations of accidents and thus be able to produce pin maps automatically, based on selective subsets of the accident file; e.g. all night-time fatal accidents.

b. Police Departments

Police departments include all those law enforcement agencies that have jurisdiction in a county, including state police, county sheriff, township constable, and city police. Police agencies perform short term, temporary changes in the traffic system. Selective enforcement in both time and space is provided. Data is required for optimum manpower deployment and for clearer pictures of local accident problems (e.g., Is it those out-of-towners that cause all the problems?). In level II counties police agencies will be provided with accident summaries that display exceptions, not a rehash of the last summary; and tree diagrams that break accidents into the first, second, third, etc., factors that best explain accident severity in the county. Level III counties can use the automatically prepared pin maps to note selected distributions; e.g. night accidents in which alcohol violations were a factor.

c. Courts

Judges and prosecutors associated with circuit courts and local traffic courts have need for quick knowledge of a particular offender's previous violations and convictions in order to better determine sanctions. Other planning data could also be made available including the number of convictions per violation (useful in checking up on easy judges) and the relationship of violation type to accident causation and severity (useful in establishing a meaningful bail schedule).

d. Medical

The medical users include administrators of hospitals and emergency medical services, and county coroners or medical examiners. The latter generally require individual accident report copies for cause of death documentation. The level III counties will encode the hospital from the police report, thus giving medical administrators some idea of distribution of injury-producing accident locations and hospitals where victims are taken.

e. Driver Education

Driver education administrators and instructors will be able to review summary accident data for their jurisdiction. This information will be useful in tailoring course content to emphasize problems unique to the area. Instructors could also present summary statistics of local accidents that students could relate to.

f. Safety Councils

Traffic safety councils, commissions and associations come under a variety of names. They can play a vital role in promoting traffice safety. Given sufficient documentation on local traffic problems, they can be effective in changing priorities. For example, if a local judge is indeed lax in handing down convictions for drunk driving, the safety council can promote a change in policy with the aid of the judge's conviction/violation record.

g. New Media

Local newspapers, radio and TV can be very effective in provoking change, especially by their approach in reporting spectacular accidents. The USA seems to thrive on reactions to crises. Small crises could be documented by the media from traffic records files. Published comparisons or rankings of accident rates among counties may also help change priorities.

h. Legislators

City councils and county commissions are fairly far removed from traffic records systems and probably receive most of the information from personal experience, safety councils and the news media. While distant from the traffic records system they play a signficant role is setting traffic safety priorities. Their support is a prerequisite for any successful highway safety countermeasure program.

3.4 Location of Experimental Flow Demonstration Program

3.4.1 Program Site Selection

Three criteria will be used for site selection:

1. Need
2. Size
3. Hospitality

The need for a data program must exist. If there is not recognizable problem, no utility will result from the program. On the other hand the selection of a completely virgin site would not permit a timely completion of the program. Thus, the paradigm for selection would have a sophisticated computer-based real-time traffic records system but no existing data utilization activities.

Size of the site must be large enough to insure the validity of the evaluation results, but small enough to be within the bounds of practicality. Clearly the data flow program must be based on the state level (with the selective involvement of local jurisdictions) because existing traffic record systems are centralized at the state level.

Hospitality will obviously be dependent on social-political climate as well as technical capabilities relative to the countermeasure goals. Clearly the political jurisdiction must want to implement a data flow program for it to be successful. Technical capabilities in the sophistication of the existing traffic records system must be considered: Are data files well structured, inter-linked, accurate, automated, and timely? What is the accident reporting level (some states only reach the 30% level) and how badly fragmented (split into different files, portions not coded) is the accident data?

Besides the technical sophistication of a state's record system, consideration will be given to the competence of personnel, particularly in the area of evaluation techniques.

A casual review would restrict the candidates to fifteen or sixteen states. The states of Ohio, New York, California, Pennsylvania, Illinois, Indiana and Michigan are among the likely candidates. A final determination should be made according to the characteristics described above. The State of Michigan has been chosen for discussion purposes. Michigan more than adequately meets the requirements of need, size and hospitality, and it has performed responsively to the letter of the standards in the information handling area. A description of the Michigan traffic record system follows in the next section.

The selection of a state resolves the question of where the Information Flow countermeasure program is to be implemented. As indicated earlier there are three levels of service to be conducted: Paltry (control), progressive, and plush, and local users are to be the target group for the services. As part of the experimental design we need to select individual cities and counties. The same three selection criteria apply: Need, size and hospitality.

The Michigan State Police prepare a monthly list of cities ranked by casualty accidents per thousand population. The top 22 cities listed in Table 3-2 are jurisdictions that need assistance. All have something to improve, and hopefully it can be identified by analysis of the available traffic records.

To control for population size, the cities are grouped into four categories: Less than 25, 25-50, 50-100, and more than 100 thousand population. Then three cities from each group are randomly selected-one for each level of service. This results in the program site selection of Table 3-3 if hospitality is disregarded for the moment. The same procedure could be followed for counties as well or the counties which include the (bad) cities could be selected.

TABLE 3-3 EXAMPLE OF SELECTED MICHIGAN CITIES

	Level		
Population	I (Paltry)	II (Progressive)	III (Plush)
Less than 25,000	Monroe	Melvindale	Benton Harbor
25 to 50,000	Highland Park	Troy	Battle Creek
50-100,000	Dearborn Heights	Saginaw	Southfield
Over 100,000	Dearborn	Traverse City	Livonia

TABLE 3-2 HIGH CASUALTY CITIES IN MICHIGAN

	Casualty Accidents Per thousand Population	Total Population (thousand) Under 25	25-50	50-100	100+
Southfield	18.6			X	
Hazel Park	18.4	X			
Dearborn	18.1				X
Pontiac	17.2			X	
Highland Park	16.4		X		
Saginaw	16.2			X	
Walker	16.2	X			
Detroit	15.7				X
Benton Harbor	15.6	X			
Muskegon	15.5		X		
Troy	15.3		X		
Southgate	14.9		X		
Lansing	14.2				X
Traverse City	14.1				X
Fraser	13.9	X			
Roseville	13.6			X	
Melvindale	13.6	X			
Dearborn Heights	13.3			X	
Battle Creek	13.2		X		
Livonia	12.7				X
Escanaba	12.3	X			
Monroe	12.3	X			
		7	5	5	5

3.4.2 Existing Traffic Records System in Michigan

The State Police and Department of State are the two organizations involved in the Michigan traffic records system. The Department of State accumulates driver and vehicle records, and the State Police accumulate accident records. A summary of each accident report is filed in the appropriate driver's record. A copy of the computerized accident file is supplied to the Highway Department with accident locations on state trunklines already coded. Figure 3-4 shows how the information flows.

Michigan State Police

The Michigan State Police headquarters provide for the accumulation and utilization of police accident reports from all jurisdictions in the state. Two file maintenance procedures are performed. Data elements for generating statistical summaries are coded onto mark sense sheets and accumulated in a magnetic tape file. The locators of accidents on state trunkline roads are coded by control section positions and included in the tape file. Copies of the tape file are supplied to the highway department for use in their spot improvement program.

The second procedure involves microfilming the hard copy accident report. The microfilm contains miracodes for retrieval. Accident and driver identification information is coded onto paper tape for use in generating the microfilm miracodes. The paper tape is also supplied to the Department of State for inclusion in the driver records as part of the driver improvement program.

Besides supplying accident information to the highway department and the driver record, the state police prepare standard summary data. This is supplied to the National Safety Council and published monthly and annually. A unique service is provided to local users, 350 jurisdictions of over 2,500 population receive summary statistics of accidents in their area. Ten countywide traffic councils are also supplied with a combined summary data for all jurisdictions in their county. Regional meetings are conducted to elicit local user reactions and participation.

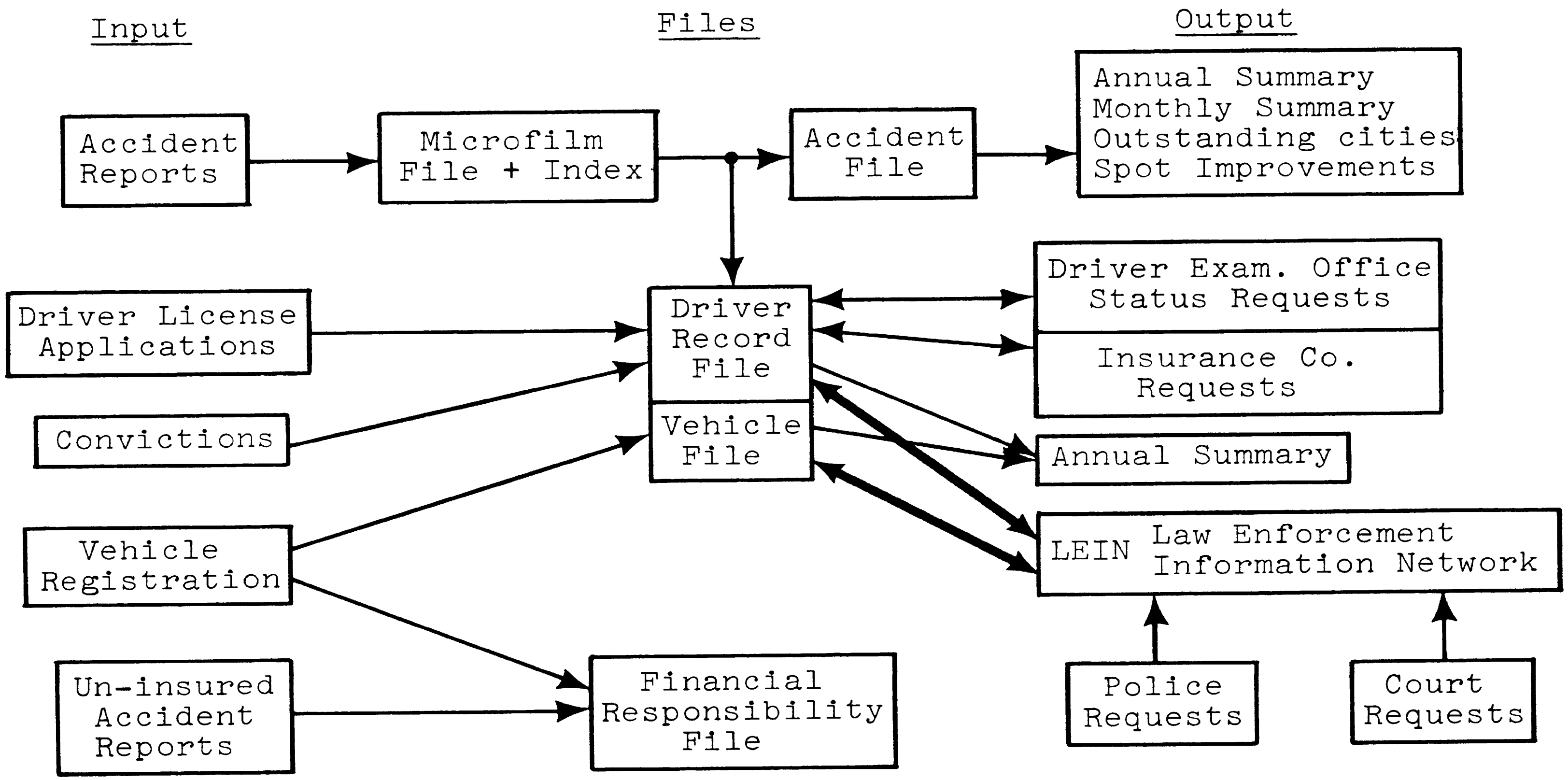


Figure (3-4) Diagram of Michigan Traffic Record System

Expansion of this program to provide specialized services, such as manpower assignment to individual areas, would be fairly straightforward.

Many state code accident locations along state trunklines (by control section, mile posts, links and nodes, etc.). The Michigan State Police are going a step further and plan to code the location of accidents on all roads in the state by distances from intersecting streets that have been coded. English street name data are read directly by the computer and given a code. Dictionaries for the street names and code values have been completed for eight pilot counties (Clinton, Ingham, Isabella, Kalamazoo, Mason, Oakland, St. Joseph, Tuscola). Plans to complete the remainder of the state are underway. This scheme will aid the identification of high accident locations. The addition of X-Y coordinate data would provide the ability to automatically plot a pin map. Dictionaries for the street names in eight counties have been completed.

Department of State

The Department of State maintains both a driver and vehicle file on-line. The driver records are updated weekly. Drivers with more than 12 points, involved in fatal accidents or several accidents in two years, are referred for re-examination or sanctions (e.g., Revocation). Paper tape from state police accident microfilm index preparation is provided for the Driver File. The information includes:

Accident Report Number
Driver Identification
Number of Vehicles, Injuries, Killed
Coded Driver, Vehicle, Roadway Factors
Violation Indicated

Plans are being made to produce an annual driver facts summary; breaking down by age and sex, license type and restrictions, number of accidents, number convictions, etc.

Courts are supplied with standard report forms for conviction records. Optical scanners record the reports. Some courts use their own forms or plan to supply data on magnetic tape. Delays up to one year are common.

Courts gain access to the driver records through the police LEIN system (Law Enforcement Information Network) as an aid in sentencing, but the system is not fully used. More terminals are planned for placement in courts, and some courts are planning their own records systems, e.g., Grand Rapids and Warren.

Insurance companies make 5,000 to 6,000 inquiries a day for five-year driver histories. Responses to the inquiries are supplied on paper tape either by mail or through license examination office terminals in slack time. Requests are batched, run at night and mailed with a 2 to 3 day response time. A $2.50 fee is charged per record requested.

Financial Responsibility (FR) information is stored in a separate file and an FR status code is recorded in the driver's record. The driver's record also includes a code for completion of a driver's education course.

Vehicle registration data is also stored on-line. The vehicle owner's driver-license number is included as a driver-vehicle file link. No vehicle-inspection or accident-damage data is stored.

There are currently 95 teletype terminals with access to driver records, primarily in driver examination offices. Five CRT terminals are used in the computer center. A new system of some 300 terminals is being installed, that will provide more terminals in examination offices, courts and external organizations. Police have access to both the driver and vehicle files through 180 terminals in the LEIN system. The LEIN system also contains warrants and stolen-property information, and it has an on-line communications link with the driver and vehicle files.

3.5 Elements of the Information Flow Demonstration Program

One state will be selected with an existing, automated traffic records system. Three different levels of information service will be provided to counties with poor accident histories. The first-level counties will continue at the same level, as a control. Services for the second level counties will be based on the state level, and will provide extensive push (dissemination) services and will offer pull (request) services. The third level counties will receive second level services plus a local traffic-records accumlation and analysis facility.

A statewide traffic records committee will be formed to guide the development of the program. An initial user survey will be conducted in the selected counties in order to determine user needs.

The utilization of accident records will be emphasized. In order to determine accident rates, an exposure survey will be conducted. Evaluation of utilization will be based on three measures: (1) Did the user know of the resource? (2) Was the product relevant and understandable, i.e. did the user look at it? (3) Did the information influence his actions? Data on usage will be collected before, during and after the program.

3.6 State Traffic-Records Committee

Chapter five of the Traffic Records Highway Safety Program Manual calls for the development of a traffic records committee to provide both administrative and technical guidance for 1) adopting the Traffic Records Standard requirements; 2) assessment of present capabilities, resources, and constraints; 3) establishing goals for the system; 4) evaluating the traffic records system; 5) reporting progress to the Governor's Representative and the NHTSA; 6) developing local and private participation, cooperation, and support; and 7) restricting record use to authorized persons and needs.

Administrative and technical responsibility in the following areas will be represented on the committee: Law Enforcement, motor vehicle inspection, driver and vehicle licensing, highway engineering, traffic engineering, driver education, legal judical/court, and public health and medical. The committee will be assisted by an interdisciplinary team of management, systems, and highway safety experts in planning and implementing a traffic records system.

The State of Michigan does not currently have a traffic records committee with the breadth of responsibilities outlined above. However, the state police have formed a committee to review requests for special bi-level reports to be collected in addition to the standard police accident report. One possibility is that the statewide traffic-records committee can be formed using the bi-level review committee as a nucleus. The committee would report to the Governor's Representative. Representatives from the relevant local agencies in levels II and III will also be appointed.

The traffic records committee should be formed prior to the beginning of the program in order to play a major role in the development of the demonstration program. During the initial stages of Phase I the committee will meet in two workshop sessions of several days each, to lay the groundwork for the development of the demonstration program. Individual members will be consulted throughout the program. Monthly or bi-monthly review meetings will be convened as necessary.

3.7 Information Use Surveys

Two requirements exist for information use surveys. First we must identify the users and determine their information requirements as a basis for Phase I program development. Second we must record the utilization of information as an evaluation measure of program level files. Three separate use surveys will be conducted: one in each phase. During Phase I, the survey will identify the user population, their information requirements and their current information system utilization. Usage counts or frequencies will be tabulated during Phase II. A follow-up user survey will be conducted in Phase III to determine how people reacted to the program.

Why study the individual user so closely? The individual learns from experience the channels through which he has the highest probability of obtaining useful information, and will tend to select those channels whenever a choice is involved. Therefore, an information service has to demonstrate by actual performance that it provides a more profitable channel. Still, the individual's judgment will prevail rather than that of management. This behavior pattern underscores the need for local participation in the utilization of traffic records.

3.7.1 Identification of User Requirements

A preliminary task in the data flow program will be to determine who the users are, what the data needs are, which needs should be fulfilled, and what the current sources and utilizations of data are. The results of this study will assist in determining a good mix of services to provide. Realistically, user studies do not provide an exacting basis for design, but the study results should provide fruitful insights to serve as guidelines for planning.

The initial effort is in defining the spectrum of users (people and organizations) within the state, by their functions and interests. Possible questions are given in Tables 3-4 and 3-5. The results will be used to get an idea of the size and nature of the 'market' for services.

When dealing with survey questions on information requirements, a respondent cannot easily define a need which he has never experienced (such as a computer-access terminal), but such a limitation can often be overcome by the opportunity to discuss the problem in some detail. A clear distinction needs to be made between user opinion and actual experience. His lack of experience limits the value of his opinion. Questions on critical incidents and recent requirements will provide more objective results. Possible questions in this are given in Table 3-6 and 3-7.

The satisfaction of all user's needs is clearly restricted by technical, socio-political, and economic complications. The value of each requirement must be determined and weighted by its relative importance in order to begin analysis of services to be provided. To simplify analysis, the number of requirements will be restricted. A preliminary list would include:

TABLE 3-4 USER IDENTIFICATION QUESTIONS

Name____________________ Age__________

Address_________________ Profession________________

Employer_________________

Title/Position_____________ Responsibilities/Duties______

Number of Persons Reporting to You____________

Which of the following categories best describes your job?

- Law Enforcement
  - Police
  - Courts
- Education
  - Driver Education
- Public Information
- News Media
- Legislation
- Driver Licensing
  - Examination
  - Improvement
- Engineering
  - Traffic Engineering
  - Highway Engineering
- Vehicle Regulation
- Medical Services
- Insurance
- Safety Council
- Other:______________________________

TABLE 3-5 USER INTEREST QUESTIONS

1. Do you have any personal or job related interest in the topic of this survey? (place an X)

/——————————/——————————/

Complete Disinterest	Passive Interest	Vital Interest

2. Have you used or had any need for traffic accident information as part of your job responsibilities in the last month?_____ in the last year? _____

3. Sometimes we have ready answers for the interviewees sample questions and could easily provide them to you. If any of your questions are urgent indicate them and if we have a ready answer, it will be sent to you.

4. (Last question on survey)

   Would you like a copy of the survey results?

   Yes

   No

   Maybe

TABLE 3-6 USER OPINION QUESTIONS
Are you satisfied with existing traffic records system?
What changes are needed? (Rank by importance)
What additional services could you use? (Rank by importance)
What questions would you like to be able to ask?
Are the traffic record summaries helpful?
Any suggestions on improved formats?

TABLE 3-7 USER REQUIREMENTS
List the last three requirements you had for traffic records information?
List three current accident data questions you have - either of personal interest or job related.
How soon would you normally like the answers to each question?

1. Delay in receiving product.
2. Portion of relevant material overlooked, in terms of time spent looking elsewhere.
3. Form of product.
4. Effort of user to communicate with system.
5. Portion of irrelevant material provided.
6. Currency of data provided.
7. Adequacy of file or sample size in providing a significant answer.

The follow-up user survey in Phase III will consider the measure of agreement between these requirements and the actual performance judged by the users following Phase II.

The final question (present sources and utilization) is really the Phase I information use question, to be considered in the next section.

3.7.2 Measures of Data Utilization

Traffic safety countermeasures (e.g., a new left turn lane) are initiated and implemented as a result of decision making. The nature of these decisions is determined by the flow (or lack of flow) of information to the decision maker. (We presume that all decision making is based on information supplied.) Thus the interface between the information system and the user/decision maker is critical.

It is possible that the information used by a decision maker might come principally from a newspaper article, or from political pressure (say by telephone) or from an analysis of accident data performed by an engineer of the state highway department. While it seems possible that more professional analysis may lead to more appropriate decisions, it is likely true that many decisions are made as the result of public pressure and are based on less than complete information. We intend to treat the information flow relationships involved here to evaluate the effectiveness of the more sophisticated processing and analysis techniques.

First, the potential user must know of the service and its availability. Knowledge of available services can be used as a measure of user acceptance, or conversely, a measure of the effectiveness of campaigns to promote use of the service. Either way, increased knowledge of available information sources is a measurable and desirable parameter of the information flow system. Possible questions for before (Phase I) and after (Phase III) user surveys as data utilization measures are listed in Table 3-8.

The degree of penetration or acceptance of information by the user can be considered on three levels:

Quality transmission
Informative/semantic understanding
Influence/affective

At the first level we determine whether the information got there and in readable form. At the second level we look at how relevant and informative was the information. Did the user understand and

Table 3-8

QUESTIONS ON KNOWLEDGE AND USE OF AVAILABLE SERVICES

Have you heard of any of the following: - a list of state traffic records service personnel? - a list of state and national summary publications? - a list demonstration services provided? Have you ever used or contacted any of the following: - a list of traffic records service personnel? - a list of push products, state and national? Please list all the sources of accident data available to you or that you have used in the last year. Include personal contacts if they aided your understanding of the accident situation.

believe it? At the third level, in what, if any, way did the information influence his actions?

Transmission quality can be measured in terms of the quality (or goodness) of the information flow to the user. The goodness of information is a function of several parameters C_i, which are weighted with factors, w_i. The w_i are dependent on the relative importance of their corresponding parameter.

$$G = \prod_{i=1}^{n} w_i C_i \quad \text{where } C_i = $$

Currency (timeliness)
Cost
Consistency
Completeness (missing information)
Correctness (misinformation)
Compatibility (mis-match)

Quantitative measures would be defined for each parameter. For example, consider compatibility as the degree of mismatch. It can be conceptualized as a function of the number of information items needed/required vs. the number provided or available. The degree of compatibility can be computed as number of items both needed and provided (logical intersection) divided by the number of items either needed or provided (logical union). Thus if a user is provided five items and he only needed three of them the degree of compatibility would be three fifths.

Even if the user receives information he must be able to understand it and correctly interpret it. Cluttered computer print-outs with untranslated code values (e.g., 2 = night) do not lend themselves to clear and easy understanding. This can be tested in the survey by a series of questions on a complex, sample accident summary table. The questions would cover all possible interpretations of the table.

In order to explore if and how traffic records information was understood and used in making decisions we may turn to the critical incident method of interviewing. When asked to recall a specific question or problem encountered within the last month, the respondent tends to avoid biasing his answers. When considering just one incident, the following types of information can be determined.

Statement of problem.
Resolution of problem, if any.
What sources of information were used.
What traffic records were used.
Any difficulties in obtaining the information.
Any delays, excessive delays.
Problems with understanding results found.
Were the items provided informative?
Did they aid in resolving the problem?
Did they affect the solution or outcome?
In what way did the information affect the outcome?

During Phase II - the operation of the information flow demonstration program - a number of unobtrusive measures will be recorded. For example:

The number of pull requests over time.

Increase in the size of the mailing list for push services.

The number of repeat users (satisfied customers).

Changes in traffic safety program budgets.

Decisions of state or local organizations to continue particular services after the demonstration.

3.8 Driving Exposure Survey

Exposure data (estimates of vehicle miles in various classes of highway travel) will be useful to local officials in the eight level II and III counties for comparisons of accident rates over time or between jurisdictions; e.g., between county and statewide rates. The exposure data can be easily collected by mail surveys of randomly selected drivers. Each completed survey questionnaire will show the distribution of vehicle miles according to key variables characterizing the driver, vehicle, road, and environment. Mean values of vehicle miles computed from the total sample will indicate how exposure varies among classes i.e., driver-vehicle-road-environment combinations.

Besides comparisons of accident and exposure information within a county, there will be great value in comparisons between exposure distributions of a county and those of the state as a whole. Thus, a statewide exposure survey will be conducted at the same time as the county surveys. Because the state survey need not be stratified by county, its sample size and cost will be the same as for each of the eight counties.

Although the accident summaries will be prepared on a monthly basis during the second year of the program, it would be quite difficult to provide monthly exposure summaires for the Level II and Level III counties. Therefore, it is proposed that the exposure surveys be conducted at the beginning of the second year, based on mileage estimates for the past 12 months. The year-one exposure results (statewide and eight counties) would provide reference points for variations in their respective accident frequencies through the months of year-two. At the end of year-two, the annual accident totals would be compared with the annual exposure totals from year-one, as a means of measuring progress in the eight counties with respect to the state as a whole. It would also be possible to compute and compare accident rates (accident frequency divided by mileage) for the state and eight counties in various classes of drivers, vehicles, roads and environments.

Survey Samples

Each of the surveys should have samples of about 2000 drivers. In the statewide survey, 2000 drivers would be selected randomly from the state driver list, and in each of the eight counties 200 would be selected from the respective county lists. With response rates of 40 percent, there will be about 800 data cases per survey. This will allow an average of 40 cases in a maximum of 20 classes. The 20 classes is probably the largest number of driver, vehicle, road, or environment combinations that officials would want to compare in one exposure analysis; and the 40 cases per class would ensure reasonable confidence limits on the mean exposure values.

Questionnaire

The exposure surveys will be conducted by mailing of questionnaires to sample drivers. The basic exposure question will be an estimate of total miles driven by the subject during the preceding calendar year. The other questions will be independent variables which provide classifications of the total mileage estimate. Based on the findings of NHTSA contract FH-11-7293 dealing with exposure data, the variables should be driver age and sex, road type, vehicle type, day vs. night, and perhaps a few others. The distributions of total miles by road type, vehicle type, or day vs. night variables may be expressed in percentages or mileage components. (See questionnaire example in Figure 3-5.

Survey Procedures

The nine sampling lists (state plus eight counties) will be developed by random selection from the state drivers list, late in year-one. At about the same time, the questionnaire will be developed and printed, along with other mailing materials. Each mailing (18,000 total) should include a letter of request, a questionnaire, and a stamped return envelope. The mailing should be sent early in the first week of year-two. Responses will return over a period of several weeks. Coding, keypunching, and building of a computer tape file should be completed in February of year-two. Exposure data could be computed and sent to the eight counties on request during the balance of year two.

Figure 3-5

Example of Survey Questionnaire

Driver Age ____

Driver Sex M ___ F ___

Estimate of total miles driven in last year _______ miles.

Percent driving in passenger cars ____%

Percent driving in other vehicles ____%

100%

Percent driving on local streets ____%

Percent driving on divided highways ____%

Percent driving on other highways ____%

Percent driving on rural roads ____%

100%

Percent driving in daylight ____%

Percent driving at night (headlights on) ____%

100%

3.9 Level II - Progressive Services from the State Level

Since the existing traffic records system is at the state level, it seems very natural that information services be provided from the state level for the local user. While the primary mode will be push (automatic dissemination) service, pull (request) services will be offered.

3.9.1 Push Services

Push services refer to those products that are automatically mailed whether the recipient wanted them or not. Though the standard monthly accident summary statistics are often repetitive and non-informative, they can be improved. Eventually, the selective and automatic distribution of meaningful analysis to local decision makers may reap the largest benefits per cost of preparation. The exact nature of the contents, frequency, and format will be determined during Phase I by the state traffic records committee, aided by results of the preliminary user survey.

There are several fruitful forms of analysis that can be conducted on a regular basis. One example would be a program that would look at all the possible interactions of the recorded accident data elements (e.g., age, road type, time of day) and rank those that best explained accident severity in a hierarchy format. An example is AID (Automatic Interaction Detector) program, described in the Road User Regulation section.

A simpler program could look through the accident data and identify counties that stood out as exceptions to expected trends, either good or bad. To improve the format for ease of understanding, the results could be displayed on a state map showing different shadings of individual counties as good, bad, or average. Such a map can conveniently summarize data that otherwise would require a long table.

3.9.2 Pull Services

Pull services are those services provided upon request. In this question and answer mode the requestor receives only the information he asked for and only when he needs it to solve an immediate problem. While this seems like a desirable situation for the local official (i.e., not being bombarded by reports he does not need at the moment) there are potential problems. The local official may not know when, where or how to ask questions. Also, the state accident information center can be caught unprepared if they don't know a particular type of question is even going to be asked.

It is proposed that the state implement a generalized statistical analysis package that performs a number of different analysis routines on any combination of data elements. This way, if the data does exist that can answer the question, the answer can be developed.

While pull services are by definition a passive operation, it is proposed that the services be aggressively followed up. Thus the state office should determine whether their response was appropriate, and whether the county would prefer another approach.
For example, if the requester asks for a two-way or cross tabulation of "time of day", and "weather conditions" for all pedestrian accidents in the county, the information center will review the results before mailing, and where possible make suggestions for further analysis.

Other requests may be less specific and require the state staff to help define a requestor's problem. For example,"The Mayor has asked me to look into the problem of pedestrian accident causes and countermeasures last week, after his son was hit by a car on the way home from school. Can you be any help?" Rather than hit him with pages of numbers, the state analysis staff could prepare a 3-D histogram of pedestrian accidents by age and time of day in order to gain an overview of the problem. Likely, young children on the way home from school would stand out as a clear hazard. (See Figure 3-6)

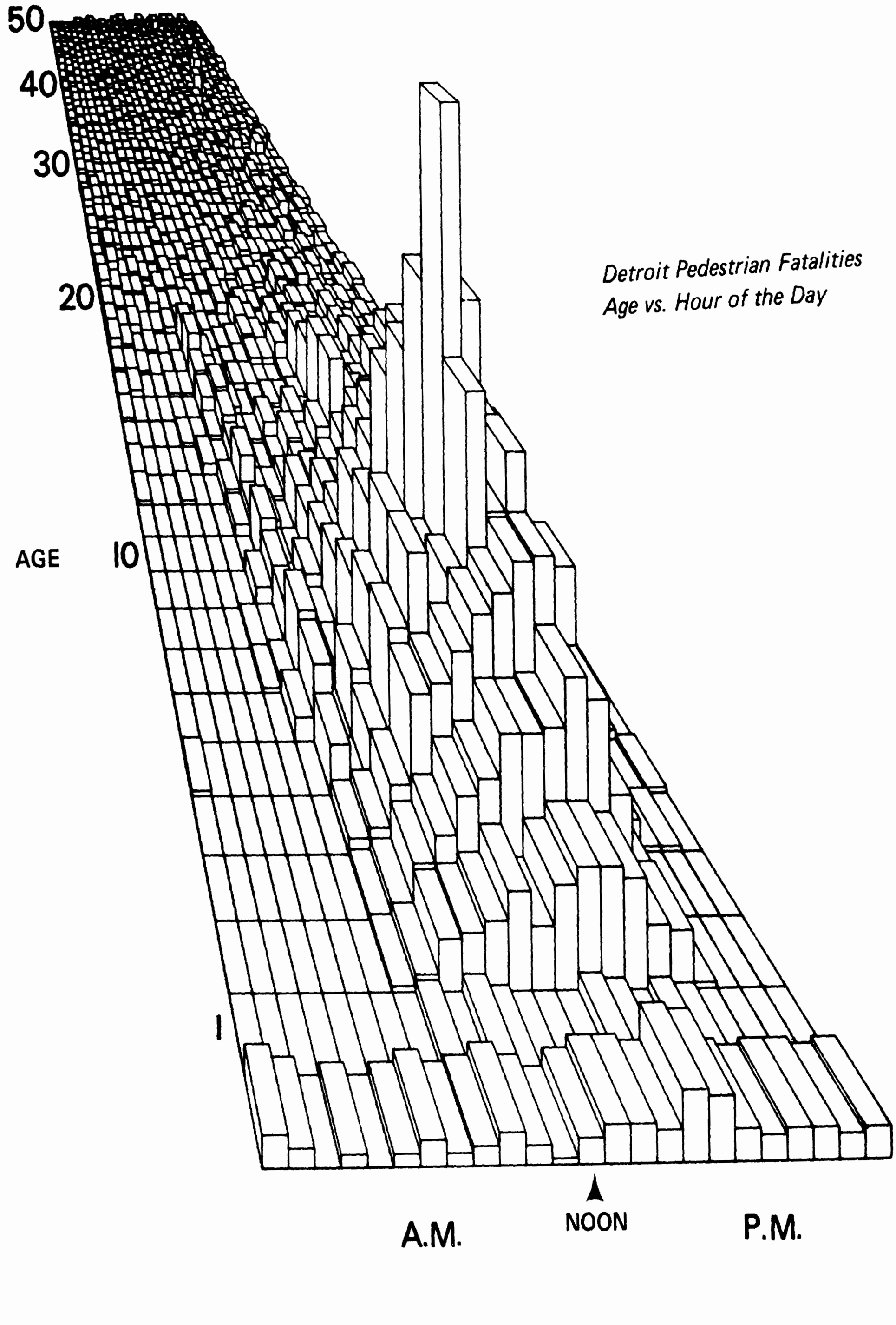


Figure (3-6) Example of Pedestrian Fatality Analysis

3.10 Level III - Plush Service at the Local Level

The four Level III counties will receive the royal treatment. All facilities available to the Level II counties will be made available as well as the establishment of local accident records system and analytical capacity.

3.10.1 Local Traffic Records Accumulation

It is proposed that a local traffic records data base be established. The responsibility for processing would remain at the local level, to encourage better beedback between file maintenance and user. If the local police department finds errors in files, the system will be improved. Local involvement and team spirit may enhance the information flow.

The accumulation of accident reports on a county-wide basis provides the opportunity to code many additional data elements not accumulated at the state level. Examples include - accident location by intersection and X, Y coordinates, and collision configuration i.e., details from the collision diagram as an aid to highway and traffic engineers. The local file can also be maintained on a more timely basis, thus being more responsive to local demands.

A central, time-shared computer could support the four individual county data bases at the state level. This arrangement increases usage of the state police computer, but leaves the responsibility for file contents and analysis at the local level. Conversely, this arrangement avoids the installation of separate computer files in each county.

3.10.2 Local Traffic Records Analysis

The Level III counties will receive a fairly aggressive treatment. The four counties will have the full-time services of the analysts who should jointly have some background involvement with the state traffic records system, statistics, law enforcement and traffic engineering. These people would work directly with local officials to interpret the data coming out of their files, and to assist them in planning countermeasures in response to the data.

The two analysts will act as consultants and will be supported by the demonstration program. However, any countermeasures planned in response to the data would not be funded by the demonstration program. Rather, it is assumed that current community budgets would be modified in response to the information presented, or that the information would be used to initiate requests for budget increases. In fact, these reactions to the data are good indirect measures of data utilization.

3.10.3 County Traffic Safety Teams

The level III counties will establish traffic safety teams, following the example of Virginia and Wisconsin. The teams will consist of at least three representatives: the county highway commission, law enforcement, and a highway safety council or driver education administrator. They will meet at least bi-monthly to review the summaries and location maps of local accidents with the last year. The team then makes recommendations to the appropriate organizations or digs for more facts in areas of interest. Beyond the utilization of traffic records, the local teams provides an opportunity for John Q. Driver to comment on dangerous situations.

3.11 Time Plan

This demonstration is a three year program consisting of three phases:

Phase I - Development and Implementation (18 months)

Phase II - Operation (12 months)

Phase III - Evaluation Analysis (6 months)

The first phase involves establishing the state traffic records committee, designing the exposure survey, conducting the initial user survey, building the analysis staff, preparing computer programs, and establishing the four Level III county accident data bases. It is expected that this phase will take over twelve months of concentrated preparation. It is important that Phase I be completed at the end of a calendar year.

Phase II of the demonstration program will be operated for one calendar year, January 1 to December 31, to coincide with the standard accident summaries. The exposure survey will be mailed the first week of January (asking about the previous year's driving experience) and completed by the end of February. Certain indirect measures of data utilization will be collected during the year.

During the third phase, evaluations will be performed on the effectiveness of Phase II operations. A follow-up user survey will be conducted to obtain direct measures of utilization and analysis will be conducted on all the utilization data collected during the program. To the extent possible, cost/benefit relationships will be determined. This phase is estimated to require six months for completion

3.12 Resource Requirements

The level of resources required for the implementation of a particular information utilization countermeasure program will be dependent upon site selection, results of the users study, and experimental design controls. The selection of a state with a relatively unsophisticated traffic records system would require more resources than a state with a sophisticated system. The results of the users' needs study may uncover a data analysis requirement not available in the existing state traffic records system. The level of resources will be controlled as part of the experimental design developed in Phase I.

The lowest level of implementation would require a state-central staff of six, consisting of a manager, one clerical/secretarial, and four staff with computer programming and statistical/data analysis experience. This staff with support funds (e.g., computer time) is estimated to cost between $120,000 and $200,000 per year.

The highest level of implementation within one state would require a central staff of sixteen (including six field representatives) and support of local or county traffic records systems. This program is estimated to cost between $400,000 and $500,000 per year, plus $100,000 for each local records system established.

Additional costs for software development are estimated at $100,000 over the first two years, and additional costs for evaluation and user studies over the duration of the program are also estimated at $100,000. The proposed exposure survey will cost about $50,000.

4.0 ROAD USER PREPARATION

Road User Preparation consists of those programs which attempt to raise the performance of road users through education and training of all types, including propaganda. This section contains four major parts:

1) A review of the priorities for countermeasure activity in all types of road user preparation programs.

2) A brief discussion of the background to child pedestrian and bicycle safety programs, and some recommendations for demonstration program activity.

3) A brief discussion of the background to safety propaganda campaigns, with some recommendations for demonstration program activity.

4) A much more detailed discussion of driver education, licensing and improvement, including the description of the comprehensive demonstration project we recommend as first priority in this area for SCOPE purposes. The content, siting, design, implementation and evaluation of the project are discussed, and tentative cost estimates given. We felt constrained to add the section "Some advice to a desperate driver educator", which offers some strategies to those who are being pressured for urgent proof of the effectiveness of programs which do not lend themselves to rapid evaluation.

4.1 Selection and Prioritization of Countermeasure Groups

In the simplest terms, a "countermeasure group" defines some set of safety system changes that may be subdivided in two ways: what you spend your funds on (e.g., training practitioners, improving management, buying equipment, etc.), and where you spend them (i.e., in which of the programs that make up Road User Preparation such as driver licensing, or driver education).

Figure 4-1 shows a two-dimensional matrix indicating these activities explicitly. Looking at the types of countermeasure effort (the "what" dimension), we identify two main categories: first, efforts directed at the design and control of programs; and second, efforts designed to give practitioners the capability of

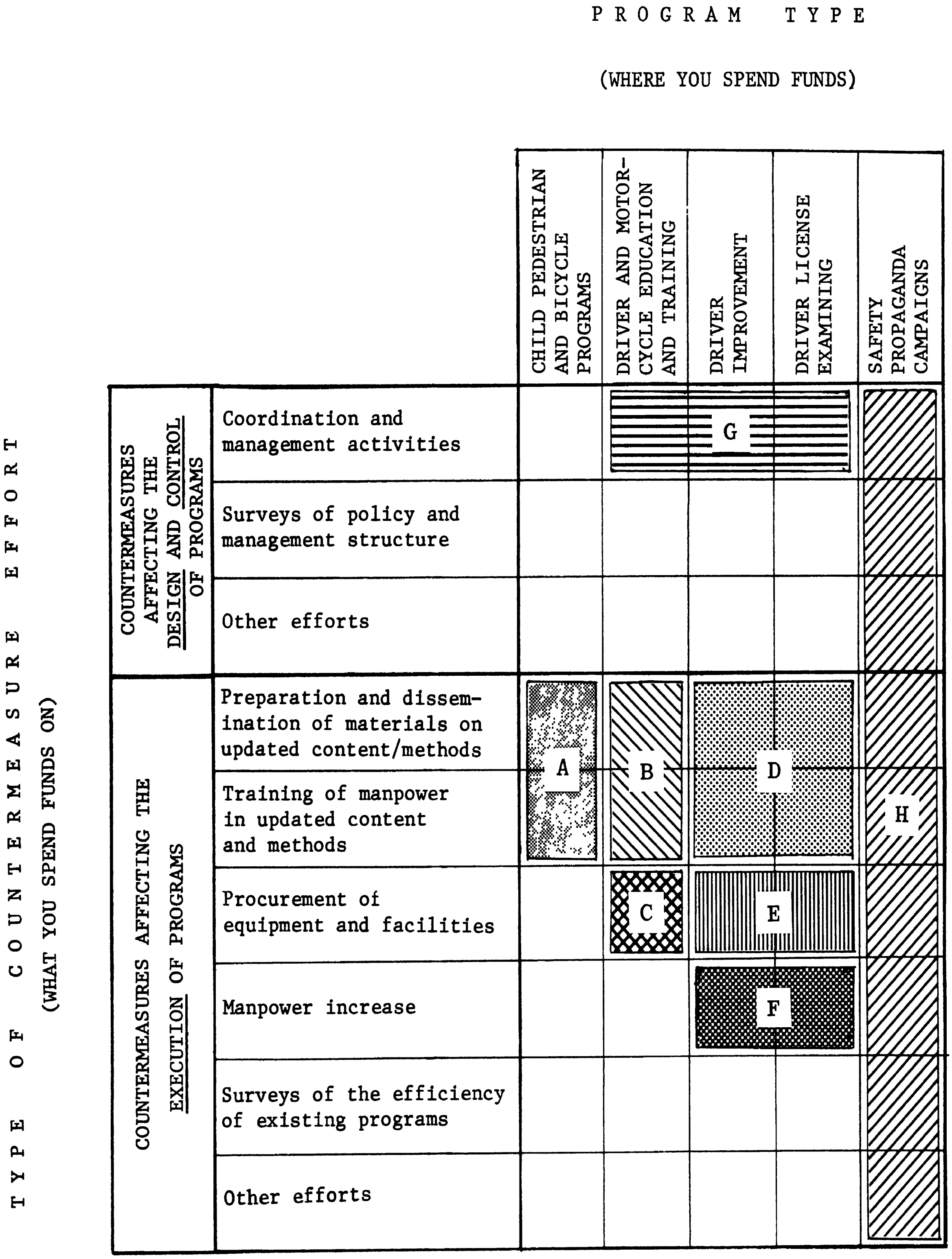


Figure 4-1: Matrix for selecting key countermeasure groups in Road User Preparation

doing a better job and/or for expanding the program. Both categories are broken down into reasonably independent efforts in the activity matrix.

The "where" dimension consists of five program types of concern to us:

1) Pedestrian and bicycle safety for children in nursery and elementary schools;

2) Driver education and training in high schools and commercial schools (including programs related to motorcycles and recreational vehicles);

3) Driver improvement activities for problem drivers and for any other experienced drivers needing or seeking individual help;

4) Driver license examining and re-examining;

5) Highway safety propaganda campaigns.

The eight selected countermeasure groups are defined in the matrix of Figure 4-1 by shaded areas. They are:

A. Child pedestrian bicycle safety - content and methods. To provide schoolteachers with updated instructional content and methods in the area of pedestrian and bicycle safety, through the development and diffusion of printed materials, and through in- and pre-service training.

B. Driver education - content and methods. To provide driver educators and instructors with updated instructional content and methods (including evaluation techniques), through the development and diffusion of printed materials, and through in- and pre-service training. This may pertain to the operation of motorcycles and recreational vehicles.

C. Driver education - equipment and facilities. To provide improved equipment and facilities for the instruction of novice motor vehicle operators.

D. Driver improvement and licensing - content and methods. To provide driver improvement practitioners and driver license examiners with updated techniques for diagnosing the difficulties of motor vehicle operators, and for ameliorating those difficulties. This to include familiarization with driver education developments, and to be done through the development and diffusion of printed materials, and through in- and pre-service training.

E. Driver improvement and licensing - equipment and facilities. To provide improved equipment and facilities for driver improvement and driver license examination activities.

F. Driver improvement and licensing - manpower increase. To provide additional manpower for driver improvement and driver license examination activities.

G. Driver education, improvement and licensing - management coordination. To coordinate the management of the parallel activities of driver education and training, driver improvement and driver license examining.

H. Public propaganda (general effort). To make a comprehensive effort to assist those who control and practice the art of public propaganda of all types in highway traffic safety.

The criteria used to define these groups include the state of the art for each of the five program types (some comment on this subject included in later sections), our assessment of achievable goals, and the information-gathering priorities in Road User Preparation. The needed information will be provided by evaluating these countermeasure groups by themselves, as well as in various combinations.

4.1.2 Priority of Countermeasure Groups and Their Combinations

Our primary reason for implementing countermeasure groups in combinations (as well as individually) is to provide for an examination of possible interaction between groups. It is likely, for example, that if manpower is increased <u>and</u> trained <u>and</u> equipment is purchased at a single site, the benefits may well be greater than the sum of the benefits from the same efforts implemented separately. Further, we must be sure to consider the interaction of countermeasure groups not just within one program type (such as driver education) but also between program types (such as between pedestrian/bicycle safety and driver education). However, if we implemented all possible combinations of the eight countermeasure groups, we would have 241 projects! In many cases we may want to try each one of these projects at different levels of effort, raising the number of possibilities to a thousand or more. However, a much smaller number of projects will serve to pose key questions on countermeasure effectiveness in Road User Preparation. Figure 4-2 is block diagram of the most useful projects. They comprise the eight countermeasure groups by themselves and 21 of the possible 241 combinations of those groups. As the diagram suggests, we recommend that as far as possible these projects should be implemented at several levels of effort.

However, we are still faced with more projects than can be reasonably implemented in a SCOPE program. Therefore we have rated the priority of each on a 1 to 5 scale as indicated in Figure 4-2. This was achieved by our judgement of each projects--

- importance as a source of information on countermeasure effectiveness
- chances of being funded
- chances of being acceptable to a site
- chances of success

In the following sections, we comment on some of the priority "2" projects as potential SCOPE programs. The priority "1" projects amount to a comprehensive driver preparation demonstration program which is described in detail. We hope that the remaining programs will find other federal support, such as "402" funds, including a substantial allotment for evaluation.

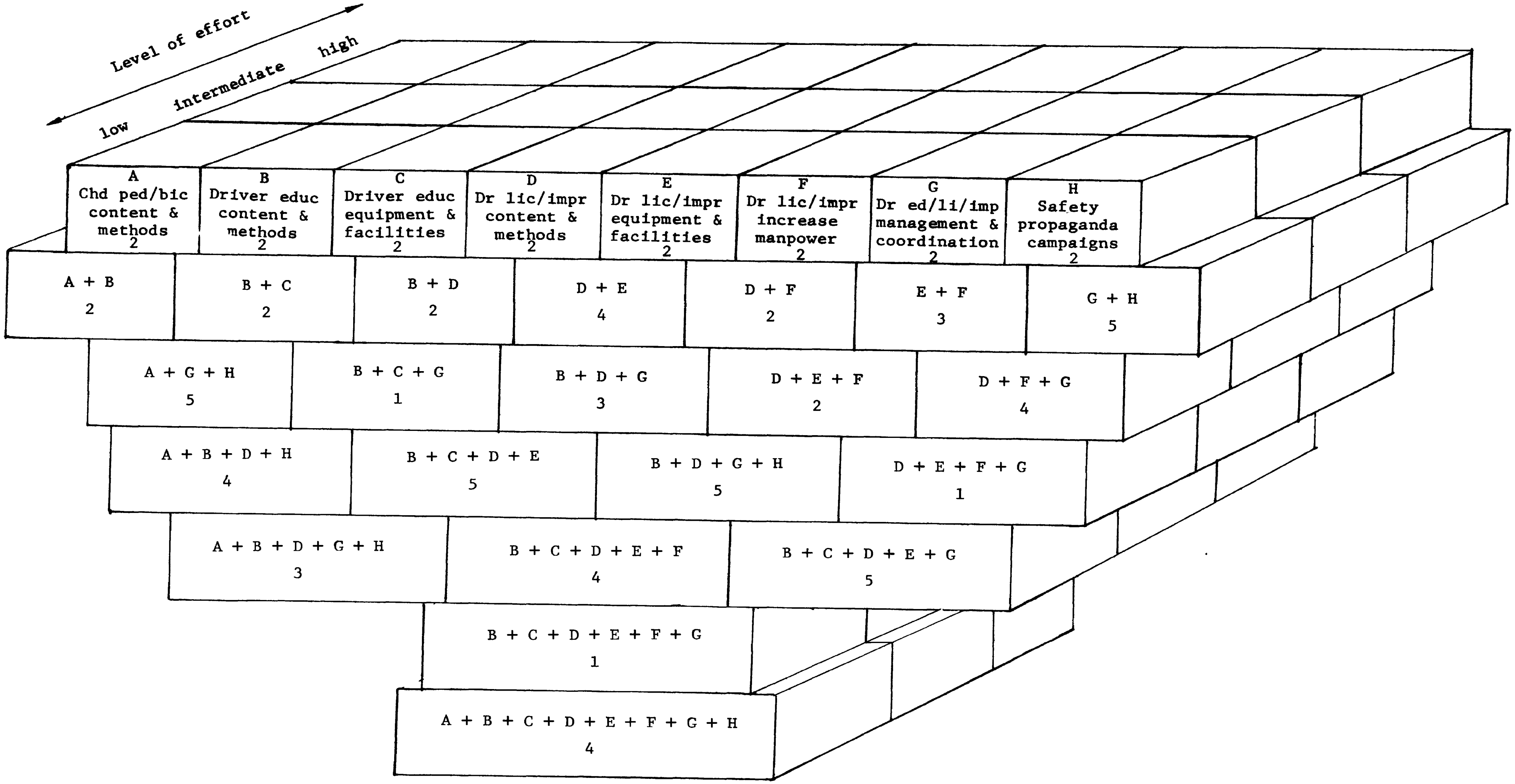


Figure 4-2: Diagram of selected combinations of countermeasure groups (numbers denote priority for implementation)

4.2 Child Pedestrian and Bicycle Programs

We have restricted our frame of reference in these programs to planned instructional activities in the elementary and nursery schools. Note that pedestrians and bicyclists of all ages are potential target groups for Traffic Safety Propaganda Campaigns.

4.2.1 Background

The schools have placed considerable emphasis on supervision of the journey to and from school. While this remains a problem in some locations, analyses of the HSRI data banks show that accidents to children under ten years old peak from between three and six o'clock in the afternoon--the period of after-school, unsupervised play. A pertinent question is whether the education and training efforts of schools can sufficiently influence the behavior of young children during spontaneous and often emotionally charged activity.

The same analyses show that six and seven year olds are relatively overinvolved in pedestrain bicycle accidents. We do not know why eight and nine year olds are less involved. We can point to the likely influence of maturation and self control; we can suggest that accident countermeasures are more effective with older children. And if both are true, then it is probable that countermeasures aimed at children under the age of eight have been inappropriate for their level of maturation. It is not unusual for those who develop or select instruction materials for the very young to evaluate them from an adult frame of reference; for example, anything in cartoon form may be assumed to be suitable for young children.

Specifically in this area, Colbourne (Reference 4-1) tested the comprehension of five, six and seven year olds of a poster typical of those produced in elementary schools; only 24% of the subjects could show any real grasp of the message, even though twice that many could read the words. In Sweden, Stina Sandels has been investigating the entire area of young children and traffic for many years (Reference 4-2). She concludes that "We cannot expect traffic maturity from children sufficient for pedestrians until the age of 9-12. Traffic education in early childhood can only lay a good foundation, but not make children perfectly safe in traffic." It is clear then that traffic education countermeasures for this age group must be: (a) complimented by adequate supervision and traffic separation efforts; and (b) much more appropriate to levels of understanding and maturity.

The elementary education system has approached traffic safety education in a number of ways. Some have embedded it in a comprehensive safety education curriculum; others have integrated it as much as possible into the other subject areas; and some have attempted to teach pedestrian and bicycle behavior specifically. Of importance to the highway safety program planner is the growing emphasis in recent years on what is called the "K through 12 approach." While the idea that traffic safety should be promoted in each and every grade is not new, there has been a marked increase in the number of states and counties who are making a formal attempt to implement an integrated K-12 traffic education currciulum. The current period is hence rather opportune for making inputs to child pedestrian and bicycle safety in the public schools.

Nursery schools, by contrast, have little organized effort in this area. They have increased in number considerably in the last few years. Churches and community groups continue to run nursery schools, but there are in addition many more on a cooperative and even a commercial basis. They are all generally licensed in matters of sanitation, fire regulations, and the like, but there is normally no administrative body accustomed to influencing their educational activities. Nevertheless, the cooperative nurseries in particular often have their own associations, and there is no reason why well formulated activities should not be promoted through them.

4.2.2 Recommendations for Implementing and Evaluating Countermeasures

For the reasons given above, we are recommending countermeasure group A in diagram 4-2 which focuses the effort on making available to teachers of young children updated curriculum content and methods, through printed materials and through in- and pre-service training. There is a great deal of misunderstanding about the effectiveness of various learning activities, particularly for the pre-school child, which this could correct.

A demonstration project of this kind would be valuable. But it has become clear in the course of this study that inadequate progress has been made in the development of updated curriculum content and methods. A small number of state departments of education are supporting some creative curriculum development work, some of it under NHTSA sponsorship; of particular note is the work of Donald LaFond* in Illinois and Maryland. An effort to bring this work together nationally (perhaps internationally) would probably provide the material for a demonstration project.

A further input is some specific recommendations for training child pedestrians which were made in a recent NHTSA sponsored study of pedestrian safety (Reference 4-3). Not surprisingly, they found that accidents to young children were frequently related to darting out into the road and dashing across intersections; they suggested that pre-school children should be taught how to play so that they are less likely to do this, and that school-age children should be taught to plan their routes in busy areas more carefully to minimize hazard.

Our recommendation is that, given the necessary improvement in the state of the art of instruction, a demonstration project should be implemented in a small number of nursery and elementary schools (perhaps three of each). They would be selected from areas with a significant amount of child pedestrian and bicycle activity in the streets. The primary project activity would be a one or two

---

* Presently Coordinator of Pedestrian Safety, Maryland State Department of Education.

day workshop for the teachers in these schools, to provide them with the updated content and learning activities. Live or video-taped laboratory sessions with children in real on-street situations would be particularly useful for this. Nursery teachers are generally not trained teachers, and would require more extensive workshop than the elementary teachers.

The effectiveness of teaching following these workshops can best be evaluated by systematic observation of the unsupervised behavior of the children before and after undergoing instruction. Sandels (Reference 4-2) used techniques of this kind; it would be possible for example, to measure the number of impulsive and potentially dangerous acts observed in a group of children over time. The main difficulty is positive identification of the group being studied, but there are a number of solutions to this: schools can be chosen in isolated neighborhoods; schools with uniforms (such as certain parochial schools) could be selected; or the children could be given a distinctive armband or badge to wear (ostensibly for some other purpose). Observations should be repeated at several intervals after the instruction is given, and the effects of repeating the instruction could also be tested.

Finally, to disseminate the findings and promote wider adoption of the instructional activities, narrative style case-studies should be written about the experimental school programs for publication in magazines and pamphlets.

4.3 Highway Safety Propaganda Campaigns

This section is concerned with attempts to use propaganda techniques to change road user behavior.

4.3.1 Background

Until the early 60's, the effectiveness of safety propaganda went almost unquestioned. Then in 1963, probably the biggest single investor in safety propaganda, the National Safety Council, sponsored a review of the literature and a national symposium under Harold A. Mendelsohn with the intention of applying modern communications theory to the safety field (Reference 4-4). This did much to expose the futility of vague messages such as "drive safely".

Since that time, a number of attempts have been made (including one by Mendelsohn) to apply scientific principles to the traffic safety area, but their results have been less than spectacular. Many "successful" campaigns have lacked adequate experimental controls, or have made questionable generalizations from laboratory studies. However, the authors of two comparatively recent reviews of these studies, Haskins (1969, Reference 4-5) and Fleischer (1970, Reference 4-6) concur that there seem to be some indications that propaganda is successful in changing behavior.

Highway safety propaganda is currently under investigation by an OECD group under Gerrett Wilde, who recently characterized successful efforts as those which specify the undesirable behavior and an achievable alternative, the media, source credibility, timing and frequency, such that the message is as immediate as possible (Reference 4-7). Immediacy can be viewed not only in terms of closeness in time, but also as closeness to a physical location, or closeness in an attitudinal sense to the message or communicator. Therefore, broadcast propaganda specifying the correct procedure for negotiating a local intersection where many drivers make hazardous errors, for example, would have a good chance of success; and in fact one of the more promising campaigns did just that in Lexington--Fayette County, Kentucky (Reference 4-8).

The consensus of the recent literature is that more studies of effectiveness should be made, using adequate controls and pre-test post-test designs.

4.3.2 Recommendations for Implementing Countermeasures

One countermeasure group was identified for the propaganda area: it is "G" in Figure 4-1, which suggests a comprehensive effort rather than focusing on specific activities such as manpower training, or the purchase of equipment.

We envisage this recommendation being carried out at the local rather than the state or national level, because it seems that the most promising campaigns, and plans for campaigns, are community centered. The activity we suggest is the setting up of machinery to coordinate all persons, media and agencies which could usefully be involved in not one, but a series of, propaganda campaigns, in a number of communities which are similar but isolated from each other.

This strategy amounts to an extension of the experimental model suggested by Fleischer's study; he recommends the use of three such sites, including one as a control, to examine in considerable detail the cost effectiveness of a safety belt campaign, We suggest that several types of campaign, especially those attempting to effect (measurable) changes in behavior at specific highway locations, should be implemented at intervals over several years. Also, an important principle for good communication is to use as many different media as possible, including informal communication, and this raises many interesting issues of cost-effectiveness. To try to answer some of these, the machinery and the resulting extensiveness of the propaganda efforts should be varied in intensity from site to site. Campaigns with a similar content could then be run simultaneously, and comparisons made, giving due attention to differences in population and environmental characteristics between the communities. It is essential that changes both in attitude and behavior are measured.

A detailed description of an alternative demonstration project in the propaganda area is included in the Road User Regulation section of this Volume.

4.4 Driver Preparation Programs

This section contains three parts: first, a brief review of the very considerable difficulties which have been experienced in the past when implementing and evaluating improvements to driver education, licensing and improvement; second, a description of the comprehensive, experimental driver preparation project which we are suggesting as a SCOPE demonstration; and finally a brief discussion of strategy for those who are forced by circumstances to make urgent appraisals of programs which characteristically require rather long-term evaluation.

A keynote for this section is to be found in B.J. Campbell's commentary on the well known Goldstein-McGuire debate in 1969 (Reference 4-9):

> "In areas such as driver licensing, driver improvement, and driver education the field has too long been content with programs whose effectiveness is being questioned more and more stringently. Stern questions should be raised to stimulate an urgent search for program variations that can produce detectable benefits."

4.4.1 Background

We present first a review of the difficulties involved in implementing and evaluating road user preparation system improvements.

4.4.1.1 Background in Driver Education Activities

Driver education continues to attract a good deal of controversy. It made educational history in the nineteen fifties and sixties with its extraordinary rate of diffusion to the nation's high schools. In the National Highway Safety Program Standards, it is explicitly required as an offering for all youth of licensing age. A considerable amount of 402 funding continues to assist the states to work towards this standard. By and large, the growth of driver education has been supported by state Departments of Education, even though its inertia can no longer be maintained with the substantial claims of accident reduction found in early driver education studies.

Having failed either to prove or to disprove the basic worth of driver education at the end of the nineteen sixties, its proponents and critics turned more to consider how to improve the quality of a school offering which, despite the controversy, seems to have estab-lsihed itself. This was a healthy development, because it encouraged certain curriculum specialists to look beyond simplistic assumptions about the relationship between the number of accidents a driver has, and how good a traffic citizen he is. Instead, these specialists began to examine the nature of the driving task with the intention of identifying those behaviors which are critically important. This approach coincided with a general trend in education towards more specific definition of instructional objectives, and together they gave rise to considerable curriculum renewal which is still in progress. The first major attempt to consolidate this process nationally, was sponsored by the Automotive Safety Foundation, who assembled a team of driver educators, headed by Richard W. Bishop in order to produce a "resource curriculum" (Reference 4-10).

Meanwhile, the National Highway Traffic Safety Administration had (in addition to its support of state and local programs) awarded four concurrent "403" contracts to research organizations to draw up plans for evaluating driver education and training. Their recommendations were integrated into a long and short term evaluation plan by Harry H. Harman in a fifth contract (Reference 4-11). The key recommendation of these studies was for a driver education task analysis project which would identify in considerable detail the behaviors which are essential to driving, to assess their relative criticality, and to translate them into instructional objectives. The plans also called for considerable research thereafter to develop instruments which (a) could determine whether driver education courses fulfilled these objectives by examining the proficiency of individual students, and (b) could link those measures of proficiency with real world performance, including accident and violation involvement. This work is still in progress at this time; therefore the issue of whether driver education is effective remains unanswered. In fact, the task analysis (Reference 4-12), certain other NHTSA contracts, and the curriculum revival process in general are meanwhile beginning to influence the pattern of driver education in this country enough that we may never know whether it was as ineffective (or as harmful) as some of its most vociferous critics suggested it was in the mid-sixties.

But we should not conclude from all of this that we can conduct experiments to compare various ways of improving programs without being pressured to prove that it is worth continuing driver education at all. At the time of writing, priorities in the funding of all types of educational programs are being reassessed.in many parts of the country. Driver education is under renewed attach in some areas, along with athletics and other "non-3R" activities. Thus, many school districts cannot afford to wait until the final evidence on the effectiveness of driver education is in; they need to be able to apply better measurement to their existing programs, and to be able if possible to quote successful demonstration programs.

To-date, very little attempt has been made to evaluate the impact of the considerable investments made to improve state and local programs. For example, up to September 30, 1970, $12,172,000 of federal "402" funds were spent on the acquisition of driving ranges, simulators and instructional media.* While admittedly this sum is not surprising in view of the fact that such procurements are specifically recommended in the Standard, some educators have understandably become a little cautious in their support for a program which has become associated with large amounts of special funding.

Thus we are faced on one hand with the fact that adequate evaluation of driver education requires better long term experimentation than has been achieved to-date, and on the other with the need to assist driver educators who, because of the complexity of measure-

---

* This amount was derived from our survey of 402 programs. See Appendix A of Volume 2 for details of the 402 project file.

ment techniques, are discouraged from doing any systematic evaluation at all.

4.4.1.2 Background in Driver Improvement Activities

The Driver Education and the Driver Licensing Standards require driver improvement programs for the general public and for problem drivers, respectively.

The fact that over a million drivers have undergone the National Safety Council's Defensive Driving Course (DDC) - many of them without being asked to do so by their employers - illustrates the need for programs for the general public. The accident and violation reduction of the DDC courses, like driver education, is still an open question (see Heimstra, Reference 4-13, and O'Day, Reference 4-14).

Driver improvement programs for problem (or "negligent") drivers have been more extensively studied. In a 1969 review of the state of the art of driver improvement research (Reference 4-15), Noel Kaestner identified a number of trends; namely that

(a) There had been little attempt to apply the science of behavior modification to driver improvement;

(b) while alternative forms of driver improvement treatment were being tried, virtaully no attempt had been made to assign drivers to them on the basis of some diagnosis of their individual problems;

(c) research had been unable to isolate benefits attributable to specific components of treatment procedures; and

(d) the emphasis in treatment and in evaluation had been on violation, rather than accident reduction.

Three years later, all four of these problems are still with us to a large degree. But administrative decisions must be made, and we need information on the best programs we know how to do now. How much experimentation does it take to get definitive answers? This is well exemplified by a recent California driver improvement study.

Marsh (Reference 4-16) looks at nine alternative treatments, including one control, one warning letter, five group treatments and two individual hearing treatments. The driving records (accidents and violations) of 15,290 subjects during the year following treatment were analyzed. After some adjustments (using analysis of covariance) for possible biases attributable to factors such as the age, sex, and previous convictions of subjects, Marsh claimed that certain treatments (notably a "group educational meeting") could prevent a small but significant number of collisions. Further, he demonstrated on the basis of his conclusions that the more successful treatments could save more dollars than it would cost to implement them throughout the state. In fact, as Marsh admits, the evidence of a significant reduction in accidents is far from conclusive. Yet note that to provide the Director of the California D.M.V. with even tenuous evidence of the effectiveness of one facet of his operation, it required a very carefully designed research program. What is more, if you are in the position that an important decision must be made regarding the adoption of a program, and the amount of accident reduction attributable to that program is very marginal (as is almost always the case with "driver" countermeasures), then very large samples of people are required if you are to place any confidence in the results.

Note also that this complex study leaves many questions unanswered. For example:

- should we attempt to evaluate in the same way other possible treatments (perceptual training, other types of individual therapy, training or control of vehicles under emergency situations, etc.?)

- can we devise a way of diagnosing a driver's problem and assigning him to the most appropriate treatment we have available (rather than randomly assigning all of the drivers to different treatments)

- if the group education meeting was successful, would individualized instruction (computer aided instruction, teaching machines, etc.) do a better job?

That these questions are not answered is not a criticism of Marsh's study. The point we would make is that given the immature state of the art of driver improvement, questions such as the above are very reasonable for those who allocate funds to ask. Put another way, can we afford to test the effectiveness on accident reduction (using of necessity large and complex studies) each of a whole range of very plausible driver improvement treatments and strategies? We think probably not until we have examined a number of the alternatives on other criteria of effectiveness, such as amount, retention, and transfer to on-street performance of what is learned in a driver improvement program.

The current NHTSA driver performance research program, notably the DETRI (American University) project to improve the driving of Coastguard recruits during their basic training at Cape May, New Jersey is addressing some of the more basic issues of content and methods.

4.4.1.3 Background to Driver License Examining

The emphasis of the Highway Safety Program Standard for Driver Licensing is uniformity in the procedures which insure that "only persons physically and mentally qualified will be licensed to operate a motor vehicle". Accordingly, a contract was let to Spindletop Research in 1967 to review state practices and to recommend a set of licensing standards which could be adopted nationally (Reference 4-17). Their conclusions were, however, that a set of uniform national standards for "screening out the physically and mentally unqualified" would not be in the best interests of improving traffic safety.

In 1963, Leon Goldstein, in a paper to the National Safety Congress (Reference 4-18), had presented a powerful argument that it was inappropriate to operate driver licensing after the number of a personnel selection process. The argument will not be reproduced here, but it led Miller and Dimling to state in the Spindletop report (Reference 4-19) that:

> "As a result of the above considerations"--(their review of Goldstein's argument and support of it from B.J. Campbell, M. Blumenthal, D. McCracken and others)--"we must conclude that the driver licensing process is, and likely will never be, an efficient screening process."

Instead of national standards, they recommended a program of research which would provide the basis for an alternative approach to licensing, namely one in which the objective was (Reference 4-20):

> "...not merely to screen applicants, but to diagnose weakness and provide assistance, counsel, or information as needed."

Their suggestions emphasize the development of diagnostic devices and remedial procedures based on a better knowledge of the driving task.

Meanwhile, in another NHSB project (Reference 4-21), the Institute for Educational Development had been examining the entire driver licensing function in order to

> "...develop plans for evaluating the effectiveness of driver licensing programs, and study sources of multi-disciplinary assistance to driver license administration."

They concluded that (Reference 4-22):

> "Current knowledge about driver proficiency and ways of influencing and controlling driver behavior is an inadequate foundation to use for evaluating driver-licensing programs. Nor will the task of improving society's control of driving behavior be achieved merely by finding and applying "standards". Instead the task is one of changing the behavior of millions of people. Doing so is a development problem--one of establishing conditions under which change will occur."

They therefore concluded that "the way to indicate the needed development is to intervene at the management level," and proposed a complex management tool ("Planning Feedback Change") similar to a PPB system, and like Spindletop, a long term program of research and development. The IED study group's definition of licensing goals are broader than Spindletops', but they agreed generally about the role of screening.

Most of the recent countermeasure effort at the state level in the area of driver examining has been towards improvement personnel had increased only marginally in a period during which the driving population has doubled. Hence evaluation of driver licensing tends to be viewed in terms of its efficiency (time taken to complete test, delay, acceptability of testing devices, etc.) rather than its effectiveness.

Altogether, driver license examining is in a somewhat similar position to driver education and driver improvement. As a matter of some urgency, it is seeking improve "content" and methods, and once again the DOT driver performance research program is attempting to provide them from some fairly basic research.

For example, Project DRIVER (Oklahoma, Reference 4-23) and Project METER (State of Washington, Reference 4-24) have evaluated several automated testing devices for their ability to increase driver knowledge through immediate feedback to an applicant on his responses, their ability to discriminate between drivers with good and poor records, and their efficiency. At this time, the Highway Safety Research Institute is constructing a bank of 2500 items for knowledge tests for various classes of motor vehicle operators (Reference 4-25); and Michigan State University is developing a research tool in the area on on-street performance (Reference 4-26).

To date, driver licensing has generally had to accept tests which are not too difficult for the majority of drivers to pass. This is because under the present screening system, failure of a test means license denial, and it is simply not feasible to deny large numbers of people driver's licenses. We have not yet had the opportunity to evaluate the adoption of a more thorough license examination, which could well be tolerated if failure meant a reasonable delay for remedial assistance rather than being refused a license.

4.4.1.4 The Evaluation Problem in Driver Preparation Programs

In Volumes I and 2, we conceptualized the traffic safety system in a causal chain fashion. Simply stated, this recognizes that countermeasure programs change in small amounts the characteristics of people, vehicles and roads; that many of these changes contribute jointly to the performance of individual people, vehicles and roads; and that these performances combine to generate the events we are ultimately interested in - accidents, violations, and conflicts, as well as normal traffic flow.

The classical problem in countermeasure evaluation is that we have tended to do one of two things:

Either a) we interpret reductions in the number of accidents and violations as evidence that the particular countermeasure we are interested in was successful;

or b) we evaluate the change brought about by the countermeasure fairly closely (such as proving that driver education students acquired a large amount of driving knowledge), and conclude from this that fewer accidents and violations will follow.

Ideally, we should like to be able to measure all programs in terms of accident and violation reduction, but we know that this is rarely possible because: (a) these are rare events, and even if we had available enough years to collect a significant amount of data on them, there would likely be some uncontrollable biases concerning which accidents and violations were reported; (b) it is unreasonable to expect any one driver countermeasure to effect a dramatic change in accident and violation behavior; and therefore (c) we can rarely be sure that any reduction that does occur is not attributable to some other factor.

What other factor should we consider in the measurement of driver preparation programs? In fact, there are a great many which can be examined at any point along the "causal chain" from the countermeasure effort, through the resultant changes in driver characteristics and driver performance, to the ultimate events. Taking them in that order, we can identify three main categories of evaluation:

1. Program evaluation: the direct evaluation of the countermeasure effort itself; includes the biographical posture of both the practitioners and the drivers in the program, as well as a wide range of measures of the immediate impact and efficiency of the program.

2. Evaluation of individual proficiency against defined objectives (including instructional objectives); this includes most of the classroom, skill and road-test type of evaluation; it spans the broad "middle ground" between countermeasure and ultimate event (hence often called "intermediate" evaluation, but note that in common usage, the measurement of all real world events except accidents and violations is also classified as "intermediate".*

3. Evaluation of individual "real world" performance through unobtrusive, direct observation of on-street behavior (e.g., using TV monitors, surveillance from following vehicle, etc.) and the judicious use of accident and violation records; this category is concerned only with the ultimate events.

The relationship between these categories and the "causal chain" is shown in Figure 4-3.

Table 4-1 is a partial list of the numerous relevant factors which we could try to measure in a driver preparation program in the three categories. Most are "dependent variables" (factors apparently affected by programs), but some are "independent variables" (factors manipulated experimentally to determine their apparent influence on programs.)

However, we do not yet have instruments capable of measuring all of these factors. A detailed appraisal of all the instruments that are available is beyond the scope of this report, but an overview is in order. It is possible with current methodology to mea-

----

* There is a statistical reason for this: it is the frequency of certain real world events - i.e., a driver typically has many conflicts and near misses before one leads to an accident.

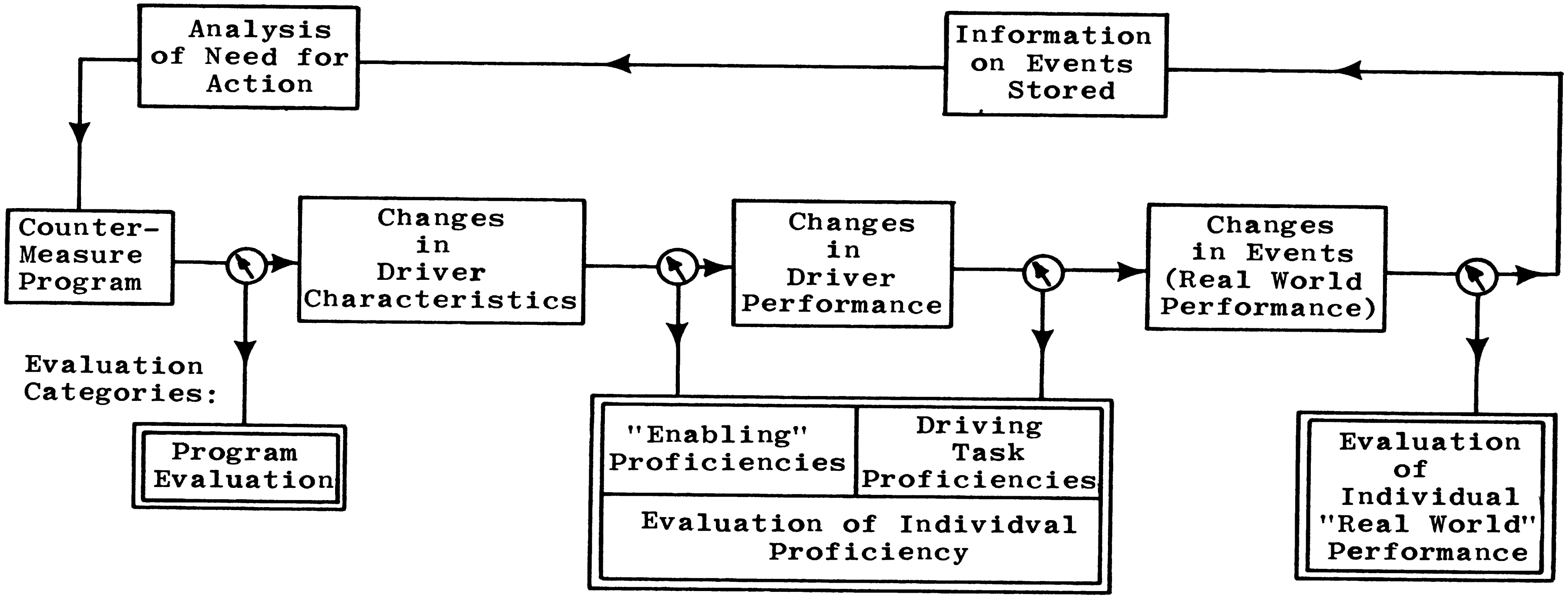


Figure 4-3: Relationship of Categories of Evaluation to the "Causal Chain"

Table 4-1

Factors Relevant to the Evaluation
of Driver Preparation Programs

	TYPE OF FACTOR	EXAMPLES OF FACTORS
1. Program Evaluation	Biographical factors	Age, sex, socio-economic level, grade point average, education level, previous driving knowledge.
	Motor vehicle related factors	Access to, and ownership of, vehicles of different age, value, type and condition; number of miles driven under different road, light and weather conditions.
	Psychological factors	Previous motivation and attitude towards a wide range of related objects, including driving; psychophysical factors (e.g., various types of reaction time); personality type.
	Practitioner (i.e., teacher, examiner, driver analyst, etc.) factors	Years of experience, qualifications, age, sex, interest, view of own role.
	Practitioner's interaction with a driver	Practitioner's view or rating of the drivers he is dealing with, and vice versa.
	Practitioner's interaction with others in different, but related, job roles	Amount of agreement with the views and methods of others doing the same job.
	Practitioner's interaction with management	Degree of communication, cooperativeness, and response to needs, as seen from both the practitioner and the management viewpoint.
	Driver's subjective evaluation of program	Acceptability, fairness, enjoyment, felt responsiveness to needs, informativeness.
	Practitioner's subjective evaluation of program	What objectives he thinks are successfully fulfilled; problems of program and their opinion of the solutions; fairness and responsiveness to drivers.

Table 4-1 continued

	TYPE OF FACTOR	EXAMPLES OF FACTORS
1. Program Evaluation continued	Actual content of program	Objectives; curriculum, treatment or test content; methods, equipment and facilities.
	Quality of actual content of program	Previous research findings on/or expert ratings of curricula, treatments, methods, equipment, and tests (including ability of tests to discriminate between desirable and undesirable driving behavior).
	Efficiency of program.	Number of drivers processed in given amounts of time and with given numbers of practitioners, delay times (especially in license examining), number of drivers who seek help outside of official improvement activities when told they have certain driving difficulties.
	Cost of program	Cost both of regular functions, and of periodic improvements, such as those suggested in this study.
2. Evaluation of individual proficiency	"Enabling" proficiencies	Driving knowledge, attitude to driving, simple and complex psychomotor skills.
	Driving task proficiency	Perception; prediction and decision-making (judgement); vehicle control (basic, and in traffic); communication to other road users - all under test conditions.

Table 4-1 continued

	TYPE OF FACTOR	EXAMPLES OF FACTORS
3. Evaluation of Individaul "real world" performance	Type of driving performed in real world	Demonstrated ability to make mature decisions about driving, especially the avoidance of certain types of exposure (e.g., driving while intoxicated or overtired).
	Frequent real world events	Incidence of unsafe or illegal acts, including conflicts and congestion, attributable to driver; frequency of adequate responses to communications from other road users and traffic control devices; frequency of adequate communication to other road users.
	Infrequent real world events	Number, type and severity of convictions and accidents.

sure most of the factors listed under the first category (program evaluation), although very few of them are in fact considered in most studies. The same cannot be said, however, for the second and third categories. In the second ("proficiency") category, there have been many attempts to measure the traditional "enabling proficiencies" of driving knowledge, attitude and psychomotor skills. In addition, there are a variety of methods of evaluating judgement and vehicle control on-street using an examiner whose presence is, of course, completely obvious to the driver. Very few instruments except certain perceptual tests, instrumented vehicles, and a very few simulators, have attempted an objective appraisal of driving task and sub-task proficiency. The third ("real world") category has even more methodological problems than the second; we have in the past been able to judge injudicious exposure and undesirable events only by analyzing accident and violation data, and for a detailed discussion of the difficulties with this, the reader is referred to that part of the Road User Regulation section of this volume which deals with the evaluation of a defensive driving course. Instruments for observing the normal pattern of (frequent) on-street events are more promising; by definition, drivers must not be aware that they are being observed, and to achieve this, human or electronic surveillance is conducted from specific highway locations or from a following vehicle; unfortunately it is almost impossible to identify and observe in this way a statistically adequate sample of drivers who have undergone a high school course or a driver improvement treatment.*

Thus, while we may hope for technological improvements in measuring real world performance, the potentially most powerful evaluation tools are not, for the time being, to be found in the third category. At the other end of the "causal chain", we can use modern survey research techniques to provide some very important information on the factors listed under "Program Evaluation" in

---

* It is possible, of course, to unobtrusively measure real world performance of known samples of certain groups of people, e.g. taxi drivers, or drivers of fleet-owned cars which have been inconspicuously instrumented.

Table 4-1, but we would still not be evaluating effectiveness on criteria which have a strong logical connection with accident reduction. The same can be said for existing techniques for measuring knowledge, attitude and skills. Therefore, what we most need are some good measures of driving task and sub-task proficiency, both for evaluating programs, and for the diagnostic testing of individual drivers. Unfortunately, we lack certain basic research tools in this area, although some attempt is being made at the time of writing to meet this need, notably the project at Michigan State University (Reference 4-26).

We have already noted that in driver education, licensing and improvement, even large and rather carefully designed experiments to evaluate the effectiveness of small parts of the driver preparation system have produced inconclusive results. The traditional reaction to this, depending upon your viewpoint, is that either the evaluation or the programs should be improved. SCOPE is intended to do both, and it appears from our review of the background and the state of evaluation of driver preparation that we are rather dependent on the success of current research efforts.

4.4.2 SCOPE Demonstration Project

This section is a description of the project we are recommending for the SCOPE program.

4.4.2.1 The Selected Demonstration Project

What alternatives are open us for a demonstration project? SCOPE projects are intended to offer some evidence of effectiveness of the "best we know" in countermeasure programs, not only to persuade other localities to adopt similar projects, but to provide a model approach for implementing and evaluating a wide range of projects in the same area of highway safety.

When seeking to identify a suitable project in driver education, licensing or improvement, we must recognize that a whole series of studies looking broadly at the evaluation of these programs have concluded in essence that:

- the state of the art of driver education and licensing improvement as countermeasures is poor enough to require extensive revision.

- a meaningful evaluation of the same programs can only take place after their objectives have been more clearly defined (and in some cases changed) on the basis of key research.

Now we could recommend as first priority a whole series of long term projects to demonstrate in an optimal fashion significant improvements respectively to driver education, driver licensing and driver improvement. But we find that any worthwhile improvement to one of these has important implications for the other two. Better coordination of all three is an essential by-product of the changes recommended by the major NHTSA contracts in this area. Therefore, we are bound to conclude that to best meet the objectives of SCOPE, the driver education and licensing/improvement functions (which together we have been calling "driver preparation") should be included in the same demonstration project. In fact, the project should set out to best the notion that it is worth the extra effort to implement the "best we know" in both these functions, rather than in just one of them.

The effort required to do this is defined by the three priority "1" combinations in Figure 4-2. The following matrix shows how these fit into four conditions which we suggest should be set up in the demonstration project:

	All drivers subject to existing licensing/improvement	All drivers subject to the "best we know" licensing/improvement
Beginning drivers receive existing driver education	No change	Do: D+E+F+G
Beginning drivers receive the "best we know" driver education	Do: B+C+G	Do: B+C+D+E+F+G

A description of each follows:

The "best we know" driver education program (B+C+G)

We anticipate that the best available program will have been defined through a project to implement a model driver education curriculum; this is a "403" program which is expected to commence during the 1971-72 school year, and which will attempt to integrate into the curriculum the relevant finding of the current NHTSA driver performance research program. It will have available to it the HumRRO task analysis, instructional objectives, and knowledge and road tests, the HSRI knowledge item bank, American Institute for Research's specifications of standards for training devices (primarily simulators and ranges), and the instruments and learning activities developed in the DETRI/U.S. Coastguard Project. Additional inputs to the design of the "best we know" program should be sought from some of the state level curriculum development projects, such as the work of Warren Quensel in Illinois.*

We are recommending that this program will be accompanied by some activity at the local management level of both driver education and licensing/improvement. The purpose of this is to familiarize licensing/improvement personnel with the changes being made to driver education.

---

* Illinois State Department of Public Instruction

The "best we know" licensing/improvement program (D+E+F+G)

Figure 4-4 diagrams the entire experimental driver preparation program, but the greater part of it concerns the suggested licensing/ improvement program. To achieve this program, we are again heavily dependent upon the outcome of the NHTSA driver performance research program. Out of the same projects listed in the previous section, we can anticipate a variety of sources for new knowledge and road tests. In addition, an integrated vision testing device is being developed by Systems Development Corporation.

The emphasis will be significantly more on diagnosis than existing license examining. We anticipate that the written portion of the standard examinations which are given to first license applicants and periodic re-examinees will be more than just knowledge tests; they will certainly require the driver to interpret and respond to a variety of traffic situations which are described or shown to them with visual aids. In addition, a more extensive examination should be assembled of items and on variables known to be of diagnostic value. This latter test should be administered to anyone scoring below pre-determined levels on the standard examinations and to all problem drivers who have been called in for administrative action.* (Under the experimental design, control groups of beginning and problem drivers and periodic re-examinees will not be required to take this). The purpose of this will be to refer initial applicants back to their driver education teacher, or to home study, as appropriate, to work on specified weaknesses. The periodic re-examinee and the problem driver will be referred to the most appropriate of a number of alternative treatments in the driver improvement program, normally to be completed on a subsequent occasion. Hopefully, the NHTSA research will have specified more closely what should be offered, but they might well include:

----

* i.e. Problem drivers who have accumulated sufficient violations to warrent action by the licensing authorities. We are not including in this project serious offenders who are subject to mandatory suspension or revocation, nor drivers who are referred through the courts to driver improvement activities, although the "diagnostic" process could be very appropriate, especially for the latter.

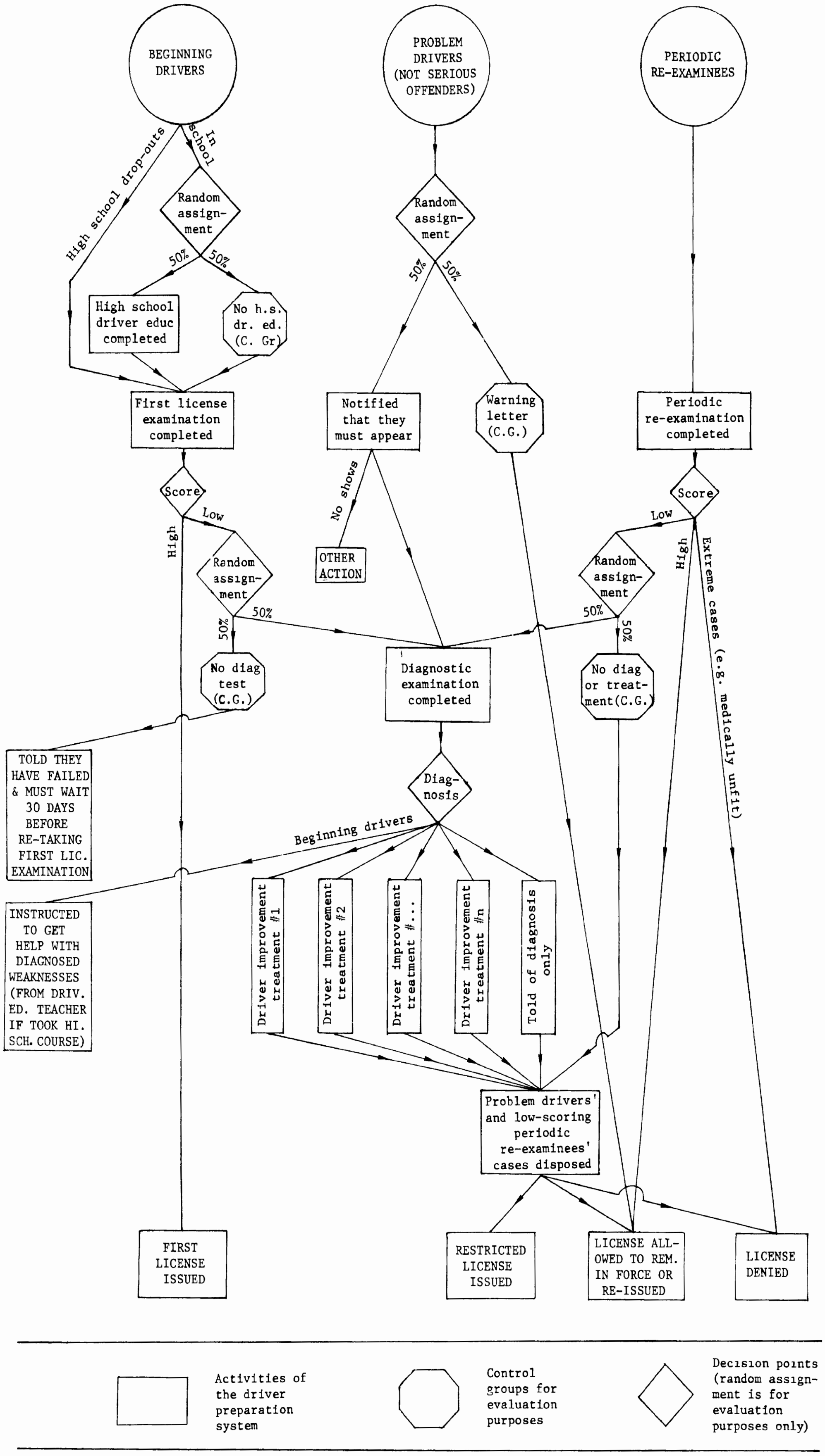


Figure 4-4: Experimental driver preparation program

(1) Individualized instruction, such as that provided by interactive computer programs (little progress at time of writing)

(2) Take home programmed instructional materials (for an excellent example, see Ontario's "Driving" (Reference 4-27)

(3) Group educational meeting (as recommended in the California study, (Reference 4-16)

(4) Manipulative skill/vehicle control and evasive maneuvers (work on this in the DETRI/Coastguard project, at UCLA, and at General Motors)

(5) Dynamic visual training (see LaFond, Reference 4-28)

(6) Individual therapy (again, see Reference 4-16)

Before each case is disposed, and the administrative decision taken to issue a license, a restricted license,* or a denial, some attempt should be made to determine whether the driver has been helped by the treatment. This will include further testing, where appropriate.

Finally, as with the driver education program, we suggest some management coordination activity; in this case it is to familiarize local driver educators with the changes in the licensing/improvement program.

The "best we know" driver preparation program (B+C+D+E+F+G)

This consists of the two previous programs operated together, with some additional management coordination activity which is made opportune through the simultaneous implementation of changes to driver education and licensing/improvement.

We expect to gain a considerable amount of valuable information if the changes are integrated in this manner. However, there is a price to be paid for complexity. In general, the larger the number of changes made and functions affected, the more difficult it is to prove the effectiveness of any one part of the effort. While this difficulty may be partially overcome with the design we are

----

* A restricted license in this context refers to one imposing restrictions on driving (e.g. to and from work only), not medical restrictions.

recommending for the program, we have to accept that it simply is not possible (even with much more strictly controlled changes) to evaluate programs for influencing driving behavior in a fully "scientific" manner.

This is not to say that we are unable to help those who must make decisions about improving highway safety programs. The evaluation approach we are suggesting for the SCOPE project is similar to that used in fields other than highway safety for projects aimed at changing human behavior. It requires two types of measures of success:

(a) something from the realm of "program evaluation" (see section 4.4.1.4) which will show whether a better process for behavioral change has been created--in this case measures of the views of pratitioners and drivers concerning the usefulness and responsiveness of the programs

and (b) a judicious choice from the available methods of appraising the behavioral change inself--in this case driving task and sub-task proficiency measures.

This is far from being a complete evaluation; but is is simply too costly to make completely thorough evaluations of the many subelements of a complex program. The strategy therefore is to look to the evaluation of this project for indications of the most promising aspects of the program, and to investigate these further.

4.4.2.2 The Selection of Sites and the Basic Demonstration Project Design

We suggest that all of the sites for the demonstration program should be from the same state. The requirements for a suitable state are that:

1. Driver education is offered in most high schools; no major (especially statewide) revisions should have been made to the curriculum during the previous three years.

2. The state should not, it at all possible, be one which has a lower licensing age for those who complete high school driver education.

3. The state should have a program for the administrative referral to driver improvement of drivers, other than those who have committed serious offenses which carry mandatory suspension or revocation. The program should require a personal appearance, and be reasonably uniform throughout the state.

4. The state should require personal appearances for license renewal and preferably for re-examination.

5. The state should have a good traffic records system.

Each site within the state consists of the "catchment area" from which a single driver licensing station draws all of the drivers it deals with - beginning and problem drivers and periodic re-examinees. The high schools involved in the project will be those within each of these catchment areas.

Driver licensing stations are generally installed on the basis of population. Thus there may only be one in rural counties, but several in a large city. We have to decide on the number and the characteristics of sites needed for the project.

Now we wish to compare situations in which:

I. Neither driver education nor licensing/improvement is changed.

II. Driver education is changed to "best we know"

III. Driver licensing/improvement is changed to the "best we know".

IV. Both driver education and licensing/improvement are changed to the "best we know".

There are two principal methods open to us to make this comparison. One is to work towards IV incrementally in a site by changing driver education one year and licensing/improvement the next, or vice versa. However, we consider it important that each change has time to "settle down," and therefore together with an initial "no change" period, this would probably require a five to six year project. The other method is preferable; it shortens the project by using more sites. In this case, situations II, III and IV are set up in three different sites in the same year, after an initial year in which the existing situation (I) is measured. We are in a better position to compare these changes if they are made in the same year, but now we need to be sure that the sites we have chosen are sufficiently similar that the comparison is meaningful. To overcome this, we need to match sites in factors which are known to be related to driving. Above all, this means examining the socio-economic character of sites; this has much to do with exposure factors (the number of miles driven under different circumstances).

The process of matching sites can become rather complex; for example, a survey of drivers could be taken in a large number of potential sites to determine their comparability on average exposure, age and income distribution, transportation needs, and other factors. However, it is probably sufficient in a demonstration program to select sites a priori from among communities with general socio-econonic characteristics in common. Social geographers have for many years made classifications of this kind using available population and economic census data.

It would seem desirable for SCOPE purposes that at least two types of sites should be chosen, so that the changes can be measured in different populations. The basic design for the suggested demonstration program is shown in Figure 4-5. This requires three each of two types of site, which we recommend should be as follows:

Type R: The catchment area of a driver licensing station in a rural county of between 25,000 and 40,000 population. We suggest as a criterion that at least 20% of the taxable property in the county should be classified as agricultural and timber. Atypical counties, such as recreational areas and those with military bases, should be avoided.

Figure 4-5 The basic design of the demonstration program

	SITE R1	SITE U1	SITE R2	SITE U2	SITE R3	SITE U3
YEAR 1	Measure existing situation only		Measure existing situation only		Measure existing situation only	
YEAR 2	Implement "best we know" driver education Begin measuring new situation		Implement "best we know" driver licensing/improvement Begin measuring new situation		Implement "best we know" driv. ed. and licensing/improvement Begin measuring new situation	
YEAR 3	Measure new situation only		Measure new situation only		Measure new situation only	

Type U: The catchment area of a driver licensing station within a commercial/industrial conurbation of between 60.000 and 250,000 population. The catchment area itself would have a population of from 50,000 to 75,000. Some attempt should be made to achieve equivalency in the three sites on the amount of suburban development included in the area. Three smaller cities of similar size, each of which has only one licensing station, might be the best choice, although again care should be taken to avoid atypical situations such as major university towns, and cities with high or fluctuating unemployment.

One of the reasons for not recommending more precise matching is that during the course of the project, important changes may occur; for example, we would like to match high schools from differing areas, but substantial amounts of bussing to achieve racial balance would render the matching unhelpful. Therefore, it is essential to gather certain background information on drivers at various times during the project in order to test our assumption that the type R sites do in fact represent populations which are consistently different from those in the type U sites.

So far, we have built into the plan a number of controls. The matching of sites and subsequently gathered data enables us to control for major population differences; and the measurement of each site under existing conditions in the first year permits us to look at each site longitudinally (before and after the changes to driver preparation). However, we still have a problem with the various instructional programs included in the project. It is that unless we withhold some of the drivers from these programs as a control group, we cannot be sure that any measured increase in proficiency is attributable to those programs.

But we cannot afford to allow the control group to select itself for reasons that are discussed in almost every recent driver education study. Fortunately, one of the few advantages driver preparation has in the area of measurement is that it involves a series of formal processes within which random assignment of individuals to different treatments is at least possible, even though the logistics may sometimes prove to be difficult. The literature on evaluation of these types of effort includes a considerable debate over the methods of controlling for variables known to be related to driver performance. In driver education studies, the controls have typically been suspect precisely because randomization has until recently been intolerable to high schools which have to explain to some of its parents that their child has been denied driver education on what must often seem very arbitrary grounds. However, the issue of effectiveness has become sufficiently important in a few parts of the country that amendments to state education regulations are under consideration to specifically permit random assignment.

Within the licensing/improvement process, there is less of a problem with random assignment to different tests and treatments. We are therefore making a number of uses of "true" randomization.

4.4.2.3 Implementing the Demonstration Project

The purpose of this section is to describe the major activities required to implement the demonstration project in each of the three pairs of sites. Measures of effectiveness will be discussed separately.

Before the Start of the Project:

We feel that the selection of a state for the project and the ensuing preliminary arrangements could best be handled through the regional offices of the NHTSA.

It will be necessary to create a small project staff in the chosen state, including at least two people with competency in research design and statistics, and at least one employee from the two state departments which, respectively, administer the driver education and driver licensing/improvement functions. Considerable cooperation will be needed from these departments by the project staff, which should preferably be independent of both departments.

Some of the project staff should be in operation well in advance of the starting date of the project to prepare the instrumentation which will be in use during year one, and to select sites in sufficient time for the high schools affected to plan for the random assignment of driver education students. Also, because of high school schedules, it will be necessary for the project to begin at the start of the school year.

Finally, it is likely that legislative approval will be required for some of the licensing/improvement changes, and the necessary arrangements should probably be initiated in this period to ensure that approval is obtained before the start of year two.

All Sites - Year 1

The activity during the first year consists of (a) preparing for the changes in year two, and (b) introducing random assignment for beginning and problem drivers to permit some measurement of existing driver preparation; (b) is the same for all sites, and is discussed here.

In all of the high schools involved (probably between two and five to every driver licensing station), the existing driver education program will be taken by half of the students. It is absolutely essential that the basis upon which students are selected to take the course is truly random, for example, by tossing a coin. As mentioned previously, there will be considerable incentives for some of the schools to assign students on some other basis; therefore, it is recommended that the project staff obtain class lists in advance, and perform the assignment themselves. A limited number of schools could be excluded from the project (if for some reason they are unable to cooperate), without causing serious experimental problems; however, it is essential to have at least one high school in each site. It is probably unreasonable to ask the schools in sites R2 and U2 (who will not be getting the new driver education) to continue random assignment beyond the first year. It is essential, however, that R1, U1, R3 and U3 schools do continue this practice at least through year 2.

A control group of problem drivers can be achieved in all the sites by selecting a portion (preferably 50%) of the drivers who would normally be called in for driver improvement activity, and sending them a warning letter instead of requiring an appearance. A warning letter is preferable to no treatment as a control because (a) it is more acceptable administratively, and (b) it is a reasonable comparison for the much more costly personal appearance treatments. We would mention again that drivers who have committed major offenses which carry mandatory suspension or revocation are excluded from this project.

Sites R1 and U1 - Year 1

Apart from setting up base-line measurement, most of year one will be devoted to preparing for the implementation of the "best we know" driver education program in year two.

What will be required, as suggested by the countermeasure group combination (B+C+G), is preparation of printed materials on the chosen curriculum, the acquisition of necessary equipment, a substantial teacher preparation effort, and some coordination activity at the management level of both driver education and driver licensing/improvement. The first two cannot be specified at this time. The teacher preparation effort should be more than a series of seminars. It should be a carefully conceived course in which teachers have the opportunity to try out new content, methods and equipment in workshop sessions, on each other and on guinea pig students. The management activity will consist of an effort to present to those in charge of licensing/improvement in the same locality an account of the changes being made in the instruction of their first license applicants.

Sites R1 and U1 - Year 2

The new driver education program is implemented, retaining true random assignment of students to the program. The existing licensing/improvement program, including random assignment of problem drivers, is continued.

Sites R1 and U1 - Year 3

The continuation of the new driver education program with existing licensing/improvement, maintaining randomized control groups in the high schools (if possible) and for problem drivers.

Sites R2 and U2 - Year 1

The effort required to prepare for the implementation of the "best we know" licensing/improvement program is not dissimilar to that for the driver education program: development of printed materials (test guides and the like), procurement of equipment, a well designed in- and pre-service training program, and in this case, the addition of manpower.

The intervention at the management level will consist of some official contact between driver licensing station manager and local driver educators in order to provide the latter with full information on the implications for their students of the new licensing/improvement procedures. Also, because these changes affect the general public, information on the changes will need to be prepared for the news media.

Sites R2 and U2 - Year 2

The implementation of the new licensing/improvement program includes the introduction of diagnostic testing and possible driver improvement treatment for low-scoring beginning drivers and periodic re-examinees. To measure the effects of this, we require a randomly selected control group from both types of low scoring drivers who will not be given the diagnostic test and treatment. As extreme cases, such as the medically unfit, are excluded from consideration (see Figure 4-4), it is to be expected that most of the drivers in the periodic re-examinee control group will receive a renewal (which they almost certainly would have received under the old re-examination system). The random assignment of driver education students could be discontinued during year 2 if it is causing the high schools difficulties.

Sites R2 and U2 - Year 3

The new licensing/improvement program together with existing driver education is continued, maintaining randomized control groups of problem drivers and re-examinees.

Sites R3 and U3 - Year 1

Preparations for implementing the entire experimental driver preparation program will be the combination of those for sites R1/U1 and R2/U2, with the exception of a more comprehensive coordination effort at the management level. This is in two stages:

1. The demonstration project director will arrange a joint meeting with top management of driver education and licensing/improvement at the state level. This will be to discuss the need for coordination of their programs to the extent that beginning drivers are able

to get the help they need if they are referred back after diagnostic testing, as well as the broader issue of each function's being fully informed about the changes in the other. These need to be dealt with at the site level, and therefore the director will seek a commitment to set up a working group to organize local joint workshops of driver education teachers and supervisors and licensing/improvement personnel in sites R3 and U3.

2. The seminar takes place fairly early in the first year of the project to give the practitioners the opportunity to participate in designing better communication between their respective programs. Problems are aired concerning beginning drivers, and an agreement is reached on the best procedure for handling them. The practitioners should also help design the public information efforts which will publicize the changes.

Sites R3 and U3 - Year 2

The combination of efforts listed under sites R1/U1 and R2/U2.

Sites R3 and U3 - Year 3

Continuation of the entire experimental driver preparation program, maintaining control groups in the high schools (if possible) and for problem drivers and re-examinees.

After the Project

It is anticipated that the analysis of the data collected will require the maintenance of the project staff full strength at least through the time of the last road tests--six months after the end of year 3.

4.4.2.4 Evaluation of the Demonstration Project

We have already discussed the fact that, given the state of the art, there is no comprehensive evaluation strategy we could suggest (even if the project were less complex) which would be immune from legitimate criticism. We have therefore selected a limited number of measures of effectiveness which we regard as the minimum to show (a) whether a better process for changing driver behavior has been created, and (b) whether changes in behavior have in fact occurred.

In Section 4.4.1.4 we reviewed three categories of evaluation for driver programs. They were: "program evaluation," "evaluation of individual proficiency," and "evaluation of real world performance." Our main recommendations, certain efficiency measures together with a survey of practitioners and drivers, and a road test, fall into the first two categories. In addition, we comment on the cautious use of accident and violation data in this project.

"Program Evaluation"

To narrow down the possibilities for evaluation is to deal explicitly with the detailed objectives of driver preparation. We would point to the truism that for a program involving numerous changes to existing practices to succeed, realistic and achievable short-term goals should be defined for those who carry out the program. Thus in this context it is important for the driver educator or the licensing/improvement practitioner to be able to say he believes his work is worthwhile, not just because it may save lives, but because he can point to things he has personally observed which indicate that the drivers he deals with are, in certain ways, better off as a result of his efforts. To try to measure all of these efforts in detail would make the project evaluation unwieldy; however, we can try to measure the extent to which the process of driver preparation has benefitted from better defined and more complimentary intermediate goals in driver education and licensing/improvement. To do this, we suggest certain efficiency measures, together with surveys of all of the practitioners involved in the project and of a sample of the drivers in the various activities. Specifications for these follow:

Efficiency Measures

Purpose: Measure workload efficiency against costs

When: Throughout entire project

Method: Normal bookkeeping and several unannounced visits annually by project staff to each program in the project

Data to be Collected: Driver education: teacher/student ratios in all phases of program; cost in each phase of program. Driver licensing/improvement: time taken to process drivers in standard and diagnostic tests; delay times in driver licensing stations; lag between assignment to driver improvement treatment and the availability of that treatment; the costs of each set of procedures, tests, and treatments.

Survey of Practitioners

Purpose: As a function of age, training, experience and interest, evaluate changes over the project period and between matched sites in practitioners' views of the realistic state of driver preparation efforts, and of the likelihood that they can make any impact on what they see as the highway accident problem. Measure the extent to which driver educators feel that licensing/improvement supports their aims, and vice versa.

When: During the last quarter of project years 1, 2 and 3.

Method: Structured interview by project staff.

Sample Selection: All of the practitioners involved in the project; this amounts to approximately 125-175 high school teachers, and 65-100 licensing/improvement personnel.

Data to be Collected:

Driver education teachers:	Driver licensing/improvement personnel:
1. Biographical data: sex; age; years of teaching experience; number of years and percentage of teaching load spent in driver education; details of preparation to teach driver education.	1. Biographical data: sex, age; years of experience in licensing/improvement; details of preparation for licensing/improvement.

2. Opinions regarding job role: reasons for teaching d.e.; their definition of the objectives it serves; the degree to which they succeed and indicators they offer of their success.

2. Opinions regarding job role: reason for doing licensing/improvement work; definition of the objectives it serves; the degree to which they succeed and indicators they offer of their success.

3. Opinions regarding program: which parts of it are the most useful and why; what are its principal problems, and how might they be solved; what needs of drivers are not met by the program.

4. Opinions regarding driver licensing/improvement: what are its principal realistic purposes; what is wrong it; how far does it succeed, and what evidence of success can be cited.

4. Opinions regarding driver driver education: what are its principal realistic purposes; what is wrong with it; how far does it succeed, and what evidence can be cited of its success.

5. Opinions regarding drivers: are most accidents caused by a small minority of drivers, and if so, what are their personal characteristics; what percentage of accidents are mainly due to driver error; which are the errors which cause the most accidents.

Survey of Drivers

Purpose: As a function of selected exposure data, age, sex, and socio-economic level, examine changes over the project period and between matched sites in certain opinions of the various groups of drivers specified in the project plan--their views of the real purposes of driver preparation efforts, and of the likelihood that the efforts can make any impact on what they see as the highway accident problem. Obtain drivers' ratings of practitioners (especially teachers) and measure the congruence they see between the purposes of driver education and licensing/improvement. Use data collected on this survey to check on the validity of site matching.

When: Beginning drivers: six months after taking first license examination.
Problem drivers and periodic re-examinees: three months after last contact with driver preparation programs.

Method: Questionnaire (largely multiple-choice) to be completed at time of road test.

Sample Selection: All drivers selected for the independent road test (see later for details of selection and likely sample sizes).

Data to be Collected: 1. Project data: Which part(s) of project driver underwent (e.g., existing driver improvement, new driver education, etc.); how they were disposed by programs (e.g., license renewed).

2. Biographical data: age; sex; socio-economic level; educational level; marital or family details. Additional items for beginning drivers on: whether in school; grade point average.

3. Exposure data: number and type of accidents and violations (to be compared with official record); type and year of car normally driven; type of roads in home area; number of miles driven on streets, rural highways and freeways, by day and by night; whether drive on job. Additional items for beginning drivers on: number of weeks driven accompanied before regular solo trips; who owns car normally used; how much of own income spent on car; principal uses of car; distance from school; how learned to drive.

4. Opinions regarding driver education: Beginning drivers: which parts were most/least useful/interesting; what are the teachers trying to do; rating of helpfulness of, and communication with, own d.e. teacher. Other drivers: what purpose is served by driver education, and should its purposes or practices be any different.

5. Opinions regarding driver licensing/improvement: what real purpose are the various procedures trying to serve; were tests fair, sensible, and comprehensible; rating of helpfulness of personnel; did they learn anything new. Additional items for drivers undergoing diagnostic examination on: was the diagnosis of weaknesses credible/helpful. Additional items for those undergoing driver improvement treatments on: how helpful was driver improvment specialist; how useful was treatment--if possible cite examples of how it has helped.

6. Opinions regarding drivers: are most accidents caused by a small minority of drivers, and if so, what are their personal characteristics; what percentage of accidents are mainly due to driver error; which are the errors that cause the most accidents.

Evaluation of Individual Proficiency

The second area of measurement we recommend is the testing of a sample of all groups of drivers on driving task proficiency. The alternative, tests of "enabling" proficiencies (knowledge, attitude and psychomotor skills) have not in the past been a sufficiently credible substitute for evidence of accident and violation reduction. We are very aware that adequate instruments for measuring task proficiency are not yet available, but we consider that the most appropriate developments are likely to be made in the area of road tests. Among relevant current research are two NHTSA contracts (Michigan State University, Reference 4-26, and HumRRO, Reference 4-12), the work of Peggy Jones at UCLA, of B.J. Campbell in North Carolina, and that of Stewart Quenault at the Road Research Laboratory (Reference 4-29).

The project staff should include the equivalent of at least three full-time people who will concentrate upon forming an independent panel of road testers. It is anticipated that they will need to spend some time before year one of the project to adapt the best available road test, to determine their reliability (consistency) as examiners, and possibly to establish standard test routes.

The specifications for the use of a road test are as follows:

Road Test

Purpose: To compare the proficiency of beginning drivers who have taken the existing and the new driver education program with those who have not. To compare the proficiency of groups of all types of drivers who have been disposed in different ways by the existing and the new licensing/improvement programs (with particular attention to differences in beginning driver proficiency between the matched sites).

When: Beginning drivers: six months after taking (for the first time) the first license examination, reached through their school or individually if not in school. Problem drivers and periodic re-examinees: three months after last contact with the licensing/improvement program, reached individually.

Method: Road test of approximately one half-hour by independent examiner from project staff.

Sample Selection: Figures 4-6 and 4-7 give a hypothetical breakdown of the three types of driver (beginning, problem, and re-examinee) into the various groups within the project. Figure 4-6 shows this for sites which have not adopted the new licensing/improvement program, and Figure 4-7 shows this for those which have. All numbers shown refer to the accumulation of drivers in one year, and are given in each case for one rural and one urban site. They are necessarily very approximate, as practices vary from state to state; the total numbers of each type of driver are, however, within the range of the workloads typically handled by driver license stations as specified in 4.4.2.2. The figures for the final disposal of first license applicants and re-examinees in the existing licensing system also reflect current practices, but the breakdown of drivers passing through the various parts of the new licensing/improvement system is, of course, speculative at this time.

The total number of drivers involved in all the sites per year will be approximately 70,000. However, the drivers of greatest interest are the beginning and problem drivers, and the re-examinees who perform poorly on re-examination. These amount to around 17,000 in project year one and around 21,000 per year in years two and three. We recommend the selection of random samples of 10% of the following groups which make up these totals:

In sites with the existing licensing/improvement program:

Beginning drivers:
1. Those granted a license at first attempt at examination.
2. Those not granted a license at first attempt at examination.

Problem drivers:
3. Those receiving a warning letter (control group).
4. Those required to appear, but did not show, and for whom other action was taken.
5. Those who appeared and underwent existing driver improvement program.

Periodic re-examinees:
6. Those who fail the existing re-examination.

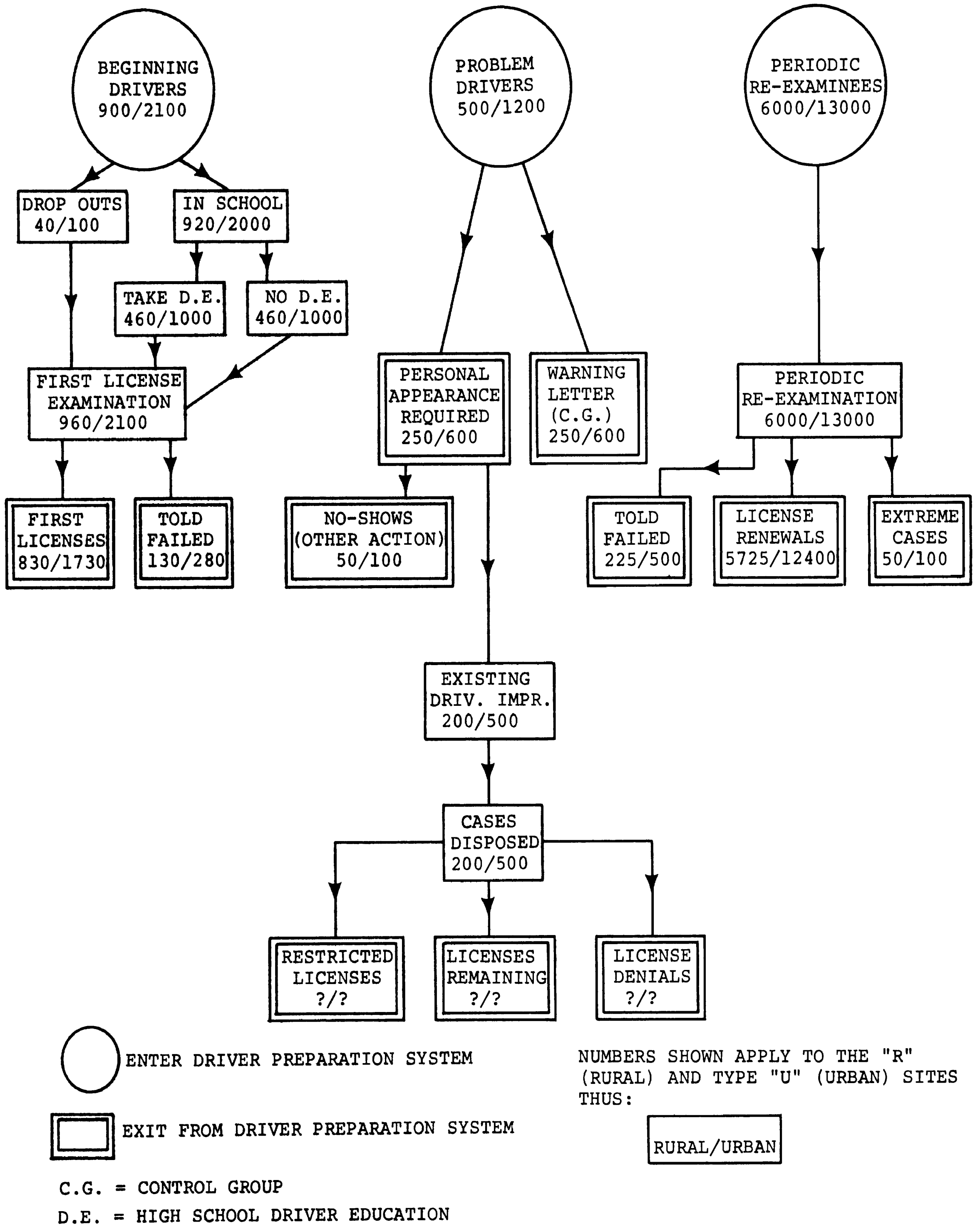


Figure 4-6: Hypothetical sizes of groups of drivers per year in sites which have not changed their driver licensing/improvement program

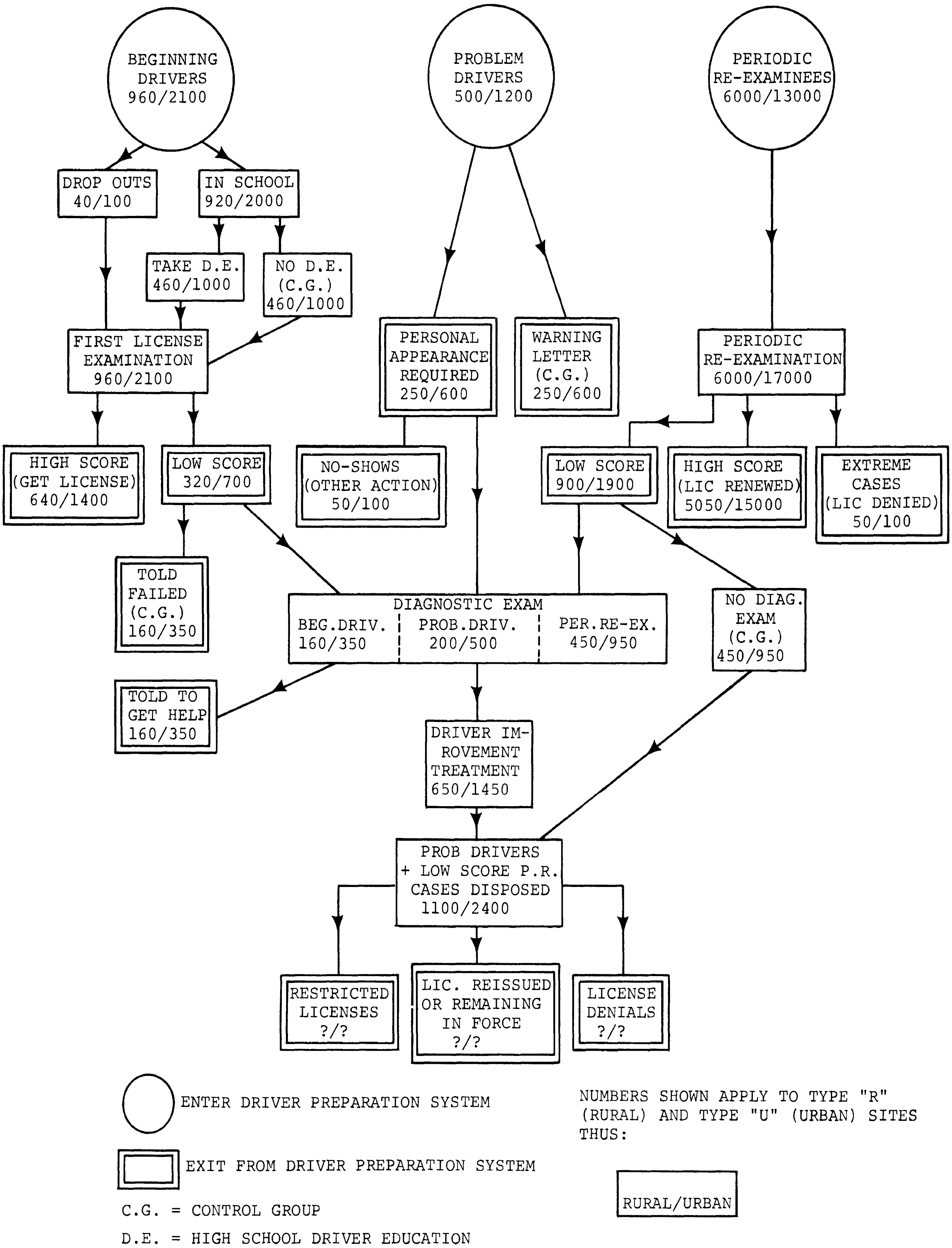


Figure 4-7: Hypothetical sizes of groups of drivers per year in sites which have implemented the new driver licensing/improvement program

In sites with the new licensing/improvement program:

Beginning drivers:
1. High scorers on first license examination.
2. Low scorers on first license examination who were told they had failed (control group).
3. Low scorers on first license examination who were given the diagnostic examination and told to get specific help.

Problem drivers:
4. Those receiving a warning letter (control group).
5. Those required to appear, but did not show, and for whom other action was taken.
6. Those who appeared and underwent the diagnostic examination and new driver improvement program.

Periodic re-examinees:
7. Low scorers on the re-examination whose cases were disposed with no further action (control group).
8. Low scorers on the re-examination who were given the diagnostic examination and driver improvement treatment if necessary.

In addition, we suggest the selection of a 1% random sample of:

High scoring periodic re-examinees

The remaining group--extreme cases among periodic re-examinees--will generally be medically unfit, and therefore are excluded from road testing.

The choices of 10% and 1% samples place road testing within reasonable cost limits. Some of the samples of the above groups will be quite small; however, they are still useable because the data from road tests (unlike the data from accident records, for example) can be used such that considerable differences are recorded among even small groups of drivers.

Note that the groups who are randomly assigned to take, or not to take, both the existing and the new driver education programs, are not specifically sampled. There are two principal reasons for this: (a) to insert a test of these groups between the end of driver education and the first license examination would discriminate unfairly against those who have been left to their own devices to learn to drive, and may reasonably take longer; and (b) it is reasonable to expect adequate numbers of those who have and who have not taken the courses to be selected in the other beginning driver samples for comparisons to be made.

Data to be Collected: Scores on driving task proficiency on-street, under conspicuous test conditions.

Evaluation of Individual "Real World" Performance

We are not recommending the use of "real world" observational techniques (car following, for example) for a project of this size for practical reasons. Accident and violation data analysis is also unsuitable for evaluating this project, except for some limited aspects, and only then providing that the traffic records of the chosen state are unusually well maintained.

Statistical reasons are given elsewhere in this volume for requiring large samples to compare programs which cannot be expected to precipitate large changes in the frequency of rare events such as accidents (see, for example, the section on Road User Regulation). Suffice it to say that the most appropriate use of accident and violation records in this project is where comparisons may be made between large groups for which there are good experimental controls. For this reason, we would exclude the use of accident and violation data as a meaningful evaluation of the differences between the various matched sites with their differing amounts of program change. The most useful comparisons on accidents/violations probably are:

1. The rates for approximately 3,000 rural and 6,000 urban area students who are randomly divided between the existing driver education program and no driver education in all of the sites together, during year one of the project. Rates should be calculated for all of the course and the no-course groups, pooled respectively; for urban and rural pools of groups; and for each of the six sites. Rates should only be calculated for individual schools if significant differences are found between the larger groups. Data will be available for all of these students over a two-year period.
2. The rates for approximately 2,000 rural and 4,000 urban students who are randomly divided between the new driver education program and no driver education in sites A1/B1 and A3/B3, during year two of the project. Rates should be calculated for all of the course and the no-course groups, pooled respectively;

for urban and rural pools of groups; for each of the four sites; and for individual schools. If considerable variance is found between and/or within schools, an attempt should be made to correlate student rating of driver education teachers with accident and violation rates. At least one year of data will be available for analysis.

3. The rates for approximately 1,500 rural and 3,600 urban problem drivers who are randomly divided between the existing driver improvement program and a warning letter treatment in all of the sites together; during year one of the project. Rates should be calculated for all of the improvement program and the warning letter groups, pooled respectively; for urban and rural pools of groups; and for each of the six sites. Two years of data will be available.

4. The rates for approximately 1,000 rural and 2,400 urban problem drivers who are randomly divided between the new diagnostic improvement program and a warning letter treatment in sites A2/B2 and A3/B3, during the second year of the project. Rates should be calculated for all of the improvement program and the warning letter groups, pooled respectively; for urban and rural pools of groups; and in each of the four sites. At least one year of data will be available.

Use of Data

Notwithstanding that the efficiency measures, the surveys, the road test, and the accident and violation data do not amount to a fully comprehensive evaluation of the demonstration project, we have suggested the measurement of a considerable number of variables; and the controls and matching we have recommended are designed to permit a variety of comparisons within and between sites. Thus it is useful at this point to draw attention again to the major assumptions which may be tested:

1. It is assumed that we are on the verge of important revisions of driver education and licensing/improvement, and that a demonstration project should display these in relation to existing programs. Hence the considerable amount of road testing of all types of drivers after undergoing the various types of instruction and examination. It is expected that significant improvement of the new programs over the old will be revealed by the road test, and hopefully in some instances, by accident and violation data.

2. It is assumed that the simultaneous implementation of the new driver education and licensing/improvement programs together with some coordination effort at the management level, will produce synergistic effects as a result of the congruence of their purposes and of their approaches to raising driver performance. Hence the matching of sites to try to minimize socio-cultural differences, and the content of the surveys. It is expected that as a result of the changes, the surveys will reveal over time important shifts in the view of the practitioner and the driver towards instruction and examining, as well as towards the accident problem. Hopefully, in the A3/B3 sites, where both driver education and licensing/improvement are to be changed, the surveys will reveal a much better climate for "official intervention" in individuals' driving behavior; in addition, it is expected that the beginning drivers, who benefit most from the coordination of driver preparation in sites A3 and B3 will show up marginally better on the road test than those in other sites, even after the complete experimental system has been in effect as little as a year.

4.4.2.5 Estimated Cost of Demonstration Project

Because the final form of both operational and evaluation activities in this project depends on the outcome of research which is uncompleted at the time of writing, cost estimates are necessarily tentative.

The following estimates are based on the implementation of the project in three rural and three urban sites with a joint annual workload of approximately 9,000 beginning drivers, 5,000 problem drivers, and 57, periodic re-examinees. Some of the costs could be off-set by charging fees for improvement treatments. Costs given are totals for three years, unless otherwise stated:

Operational Costs	Low Estimate $	High Estimate $
1. Project staff, portion of salaries attributable to coordinating operational activities, including all overheads (four years):	110,000	155,000
2. Cost of new driver education programs:		
Printed materials for teachers (curriculum guides, etc.)	1,000	3,000
Texts for driver education students	24,000	40,000
Equipment and facilities*	200,000	400,000
Management coordination activities	1,500	3,000
Project staff salaries--technical assistance, including all overheads	8,000	12,000
Project staff salaries--teacher workshops, including all overheads	20,000	33,000
Teacher workshop costs (including travel and accommodation for teachers)	8,000	14,000
3. Cost of new licensing/improvement program:		
Printed materials for practitioners (test guides, etc.)	1,000	2,000
Equipment for standard license examination*	50,000	130,000
Equipment for diagnostic examination*	30,000	60,000
Equipment for driver improvement treatments*	6,000	10,000
Take home programmed learning d.i. books	2,000	3,000
Extra operational costs (e.g. extra floor space rental at times of peak demand)	20,000	30,000
Management coordination activities	1,500	3,000
Additional manpower, diagnostic examinations including all overheads	115,000	140,000
Additional manpower, driver improvement treatments, including all overheads	135,000	160,000
Project staff salaries--technical assistance and workshops, including all overheads	7,000	13,000
Practitioner workshop costs, including travel, etc.	3,000	5,000
Total Operational Costs:	743,000	1,216,000

*Highly tentative estimates

Evaluation and Instrumentation Costs	Low Estimate $	High Estimate $
1. Project staff (including road testers), portion of salaries attributable to evaluation activities (four years)	410,000	505,000
2. Equipment	6,000	30,000
3. Printing costs	4,000	5,000
Total Equipment and Instrumentation Costs:	420,000	540,000

4.4.3 Some Advice to a Desperate Driver Educator--Meeting Short-Term Needs

This section is a brief response to the desperate driver educator, quoted in Volume 1, who will be without funds if he waits several years for some indications of the success of driver education programs. He is not alone.

An understandable, and very common, reaction to the complexity of evaluation in this area is to avoid formal evaluation altogether. Needless to say, teachers constantly evaluate student progress and the apparent success of their course out of professional experience, but they still lack tangible evidence with which to defend their efforts. The driver educator who is in this position should try to respond to two questions:

1. Whom am I trying to convince that my program is worthwhile?
2. What is available in measures of effectiveness which may help me?

Let us look briefly at some strategies which might be used in answering them.

4.4.3.1 Who Wants an Evaluation of Your Program?

We shall assume that our purpose is not the continuous process of re-evaluation which is part of curriculum development, but rather to defend an existing program, or some recent investment of resources in it. Driver educators find themselves defending their programs to a variety of people--students, parents, local and state supervisors, and occasionally legislators. For each of these, what kind of evaluation is meaningful? Some possible answers are that:

Students are primarily concerned with obtaining a license, although as in other areas of education, some are concerned about the quality of teaching, and in the relevance of what is being taught.

Parents are generally interested in being saved what is to many an unpleasant task, and in the discounts offered by insurance companies.

Local and State Supervisors usually state their interests in terms of the educational quality of programs, but when they are demanding some urgent evidence of effectiveness, it usually has something to do with funding, and under these circumstances the debate will inevitably turn to whether or not lives are being saved. Note, however, that if enough of the parents and students make it known that they want to keep driver education for the purposes noted above, many supervisors have all the evaluation they need.

Legislators are almost always concerned with financial justification, particularly as driver education enjoys special state reimbursement in many states. They most often require evaluation in terms of lives, injuries, or dollars saved. But they too are sensitive to popular demand.

Thus a program may be defended in a number of ways, depending on who it is being defended to. However, two principal courses seem open:

1. Draw attention to the hearsay support of driver education, and to the consequences (in terms of inconvenience to parents and students) of removing it from the schools, or

2. Attempt some sort of formal evaluation of the program.

Number 2 places us back where we started--facing the complexity of evaluation, but it is often the only choice.

The alternatives for formal evaluation are reviewed in detail in 4.4.2.4, and the reader's attention is particularly drawn to Figure 4-3. Unfortunately, it is easier to eliminate types of evaluation as unsuitable for short-term application than it is to retain them. We know that "program evaluation" is inadequate (we have already decided against claiming face validity and showing that it gives people what they want). We know that the time span for evaluation is far too short for the use of accident and violation data, and other "real world" measures (the observational techniques) require special expertise and instrumentation. And we know that among "individual proficiency" evaluation techniques knowledge, attitude and skill tests have been far from convincing for those who seek evidence of accident reduction. Therefore, as in the demonstration program, we are bound to look to measures of driving task proficiency.

4.4.3.2 <u>What Can You Do Now To Measure Driving Task Proficiency?</u>

There are two possibilities for measuring driving task proficiency. The first, which was suggested in the demonstration project, is to measure the whole task at one time under as realistic conditions as possible, such as in a road test in normal traffic. The second, which we recommend here, is to break the driving task down into a number of sub-tasks, and to measure as many of them as possible. To do this, we need a model of the driving task; quite a few of these exist, but most are too complex for our present purpose. A suitable model has, however, become known to most driver education teachers through the work of Bishop and the A.S.F. Resource Curriculum (Reference 4-10): it is a cycle of four sub-tasks, identify-predict-decide-execute. Being familiar with this concept, it should not be difficult for driver educators to make a case for the success of a program if he can show measured improvements in the ability of students to perform these sub-tasks. Furthermore, success in these terms has a much closer logical relationship to accident avoidance than traditional tests of knowledge, attitude and skill.

But how well can we measure proficiency on these four sub-tasks? In Figure 4-8 we have extended the IPDE model a little to consider the two-level relationship involved in driving: firstly, the interface between the driver and the vehicle, and secondly, the interface between the driver-vehicle combination and the highway traffic environment. This model defines eight possibilities for measurement, as follows:

1. Sub-task: Identify Relationship: Driver-Vehicle

   Measure: Verbal or behavioral response to critical information from vehicle

   Method: This can be checked on a driving range or on-street by observing whether driver uses the speedometer in monitoring speed, whether he has the "feel" of the road, and whether he detects salient events from the instruments (such as potential overheating, loss of charge, low oil or gas). Driver commentary can be used to supplement observations.

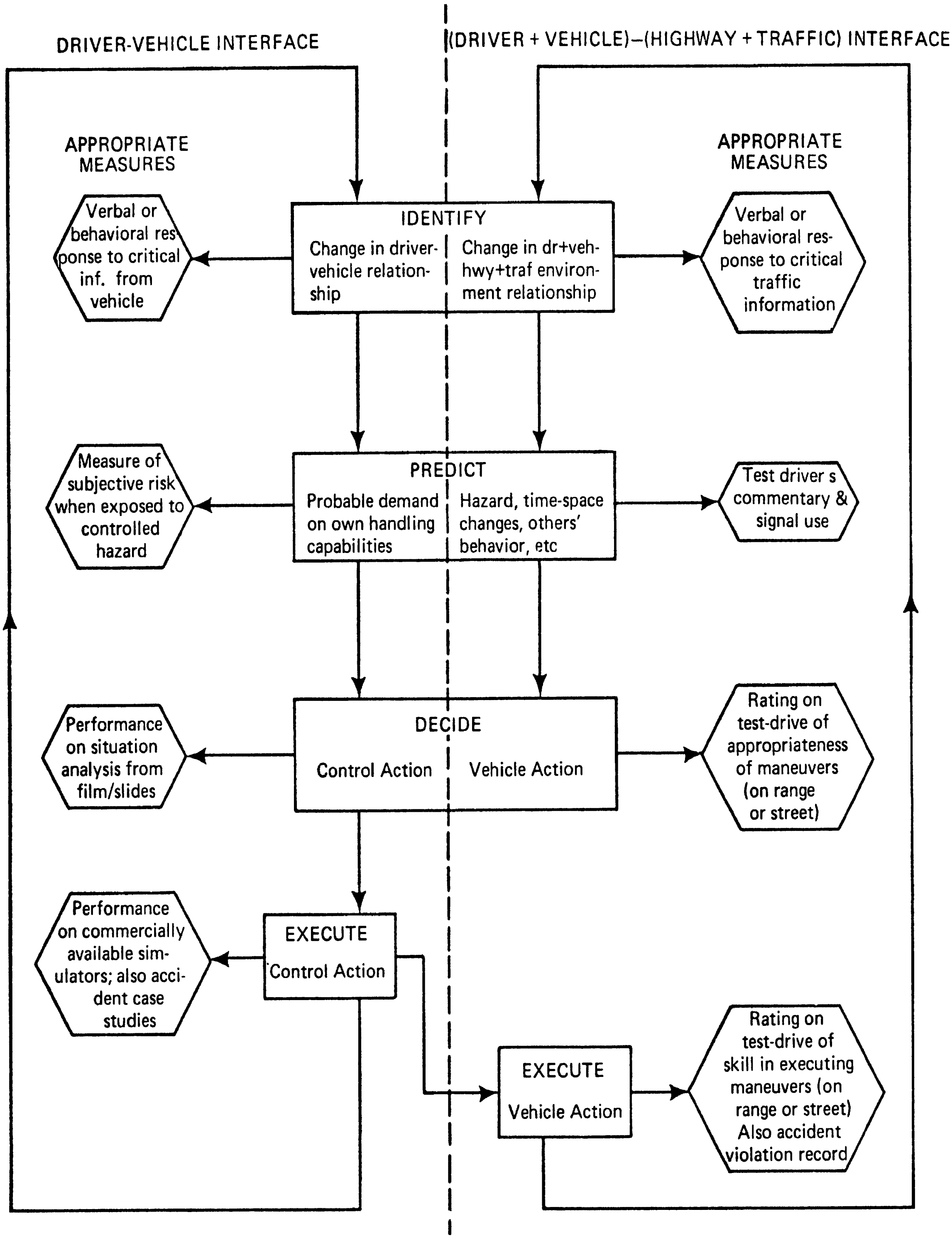


Figure 4-8: Model for generating measures of driving sub-task proficiency which can be developed without expensive instrumentation

2. Sub-task: Identify — Relationship: (Driver+Vehicle)-(Highway+Traffic)

Measure: Verbal or behavioral response to critical traffic information

Method: Use the Perception of Traffic Hazards Test developed by Kenard McPherson and Terry Cooper at Illinois State University (Reference 4-30). This can be made easily from a free filmstrip which is nationally available. Additional measurement is possible by noting the timing of "defensive" braking actions in relation to the occurrence of hazards on-street, and to lesser extent on the range and simulators.

3. Sub-task: Predict — Relationship: Driver-Vehicle

Measure: Measure of subjective risk when exposed to controlled hazard (i.e., experimentally)

Method: This is a little difficult to set up. The object is to present drivers with some hazard (using films, pictures or words), and to try to measure the amount of anxiety the hazard generates. Put another way, this is to try to measure a driver's estimate of his own ability to handle the situation. Donald Pelz of The University of Michigan has experimented with a device to record these estimates electronically, but it could be done crudely with a score sheet marked by each driver. Of particular value is the examination of situations which produce unusually low or high anxiety in a particular driver compared to his average response, as well as each driver's pattern of anxiety.

4. Sub-task: Predict — Relationship: (Driver+Vehicle)-(Highway+Traffic)

Measure: Test-drive commentary by driver and use of signals

Method: Have the driver think aloud about his predictions of the hazards which may arise from traffic situations as they develop, and from the highway environment. Observe adequacy and timing of communication to other road users, especially whether signalling is responsive rather than habitual. Supplementary measures can be obtained from analyzing traffic situations in class.

5. Sub-task: Decide   Relationship: Driver-Vehicle

Measure: Performance on situation analysis from films and slides

Method: Test the ability of drivers to select an appropriate response in terms of the control of their own vehicle to potentially hazardous or ambiguous traffic situations presented to them on slides or film. A multiple choice format, such as that used in some of the "Driver's Tests" on TV, can be used.

6. Sub-task: Decide   Relationship: (Driver+Vehicle)-(Highway+Traffic)

Measure: Rating on test-drive of appropriateness of maneuvers.

Method: Attempt to judge the driver's thinking rather than the smoothness of his driving. This can be done on-street or on a driving range. A scoring technique could be developed specifically to give positive marks for "good thinking ahead," and negative marks for tardy and inappropriate responses, and for unnecessary maneuvers. See S.W. Quenault (1968) for a discussion of unnecessary maneuvers (Reference 4-31).

7. Sub-task: Execute   Relationship: Driver-Vehicle

Measure: Performance on commercially available simulators

Method: Utilize scoring devices on simulators to determine whether drivers have tendencies to over- or under-control their vehicle.

8. Sub-task: Execute   Relationship: (Driver+Vehicle)-(Highway+Traffic)

Measure: Rating on test drive of skill in executing maneuvers (on street or range)

Method: Rate smoothness of driving, "car sympathy," and timing of control actions (especially in manual transmission vehicles).

All of the measures can be developed locally without expensive instrumentation. A selection from the suggested methods should give a driver educator a more credible evaluation of the progress of his students, especially if he attempts to measure something of each of the four sub-tasks, and if he explicates the logic of the model to his critics. Some of the tests are sufficiently enjoyable that parents could well be invited to take them in competition with their children (which may solve one justification problem). They are all included here in the hope that they may be of assistance to those who could be out of business by the time more objective measures have been developed.

REFERENCES

4-1 "Testing a poster for infants", H.V. Colbourne, Road Research Laboratory Technical Note No. 80, 1966.

4-2 "Young children's ability to understand traffic education". Stina Sandels, Proceeding of the Second Contress of the International Association for Accident and Traffic Medicine, (H.B. Wulff and E. Forsberg, eds), Stockholm, 1966.

4-3 "Pedestrian Safety, the identification of precipitating factors and possible countermeasures", Operations Reserach, Incorporated, Silver Spring, Maryland, January 1971; DOT Contract Number FH-11-7312.

4-4 "The Denver Symposium on mass communications research for safety", Murray Blumenthal, Ed., National Safety Council, Washington, D.C., 1964.

4-5 "Effects of safety communications campaigns: a review of the research evidence", Jack B. Haskins, Journal of Safety Research, Vol. 1, Number 2, June 1969, pp. 58-66.

4-6 "Cost-effectiveness of mass media communications as related to highway safety", G.A. Fleischer, U.C.L.A., Los Angeles, May 1970 DOT Contract No. FH-11-6800.

4-7 Gerrett Wilde, Department of Psychology, Queens University, Kingston, Ontario: remarks at 2nd James Lougheed Memorial Seminar, Canadian Council of Young Drivers, Kingston, Ontario, July 1971.

4-8 "An evaluation of the effectiveness of televised, locally oriented driver reeducation", James Hutchinson, Charles S. Cox, and Bennie R. Maffett, Dept. of Civil Engineering, University of Kentucky, Lexington, 1969.

4-9 Letters to the editor: "On the role of controversy", B.J. Campbell, Journal of Safety Research, December 1969, Volume 1, No. 4., pp. 146-7.

4-10 "A Resource Curriculum in Driver and Traffic Safety Education," Automotive Safety Foundation, Washington, D.C., January 1970.

4-11 "Evaluation of Driver Education and Training", Harry H. Harman, et. al., Highway Research Board, Washington, D.C., March 1969.

4-12 "Driver Education Task Analysis". Human Resources Research Organization (HumRRO) Alexandria, Virginia, DOT Contract No. FH-11-7336, 1969.

4-13 "Injury Control in Traffic Safety", Norman Heimstra, Charles C. Thomas, Springfield, Illinois, 1970.

4-14 "A before and after analysis of accidents involving students of the Defensive Driving Course", James O'Day, Accident Analysis and Prevention, Vol. 2, pp. 175-188, 1970.

4-15 "Research in driver improvement - the state of the art" Noel Kaestner, Traffic Quarterly, Vol. 22, No. 4, pp. 497-520, 1968.

4-16 "Modifying negligent driving behavior; an evaluation of selected driver improvement techniques" William C. Marsh, State of California Dept. of Motor Vehicles, Research and Statistics Section, Report Number 36, March 1971.

4-17 "Driver Licensing and Performance," Louis Miller and John A. Dimling, Spindletop Research, Vol. II, October 1968, Vols. I and III 1969.

4-18 "Driver Selection - The Lure, Logic and Logistics" Motor Transportation Transactions of the National Safety Council, Chicago, Illinois 1963.

4-19 Reference 4-17, Volume III, p. II-9.

4-20 Reference 4-17, Volume I, p. VI-43.

4-21 "Driver Licensing Program Project - Final Report" Institute for Educational Development, Newport Beach, California, 1969.

4-22 Ibid, pp. 1-2.

4-23 "Project Driver Final Report: Phase IA" Dept. of Public Safety, State of Oklahoma, The University of Oklahoma, Norman, 1970 DOT Project No. FH-11-7126.

4-24 "Project METER - An Evaluation", State of Washington, Dept. of Motor Vehicles, January 1971, DOT Project No. FH-11-6834.

4-25 Development of National Item Bank for Tests of Driving Knowledge, DOT Project No. FH-11-7616, Highway Safety Research Institute, University of Michigan, Ann Arbor, Michigan.

4-26 "Development of Driver Performance Measures," DOT Project No. FH-11-7627, Michigan State University, East Lansing, Michigan.

4-27 "Driving - a guide for improving your driving skill", Ontario Department of Transport, Queens Park, Toronto, 1970.

4-28 "The effect of dynamic visual training on manipulative driving skills", Donald L. LaFond, M.S. Thesis, Illinois State University, Normal,Illinois, August 1970.

4-29 "Convicted and non-convicted drivers - Values of drive indices" S.W. Quenault and C.F. Harvey, RRL Report LR395, Road Research Laboratories, Department of the Environment, Crowthorne, (Gives an overview of eight years of experiments) England.

4-30 "Perception of traffic hangards test" Kenard McPherson - Terry Cooper, Industrial Technology/Traffic and Safety Education, Illinois State University, Normal, Illinois.

4-31 "Development of the method of systematic observation of driver behavior" S.W. Quenault, R.R.L. Report LR213, Road Research Laboratory, Ministry of Transport, Crowthorne, England, 1968.

5.0 VEHICLE REGULATION

Three experimental programs spanning the current range of alternatives are recommended for vehicle regulation involving three levels of motor vehicle inspection: a self-inspection plan coupled with a random check lane system, annual periodic inspection as directed by State Standard 1, and a variable response diagnostic system.

The self-inspection format will require vehicle owners to certify annually the condition of their vehicles either through the vehicle registration mechanism or independently. The parallel random check lane program will detect violators of the certification requirements and will induce a continuous concern with vehicle safety quality between certification periods. Standard annual inspection will be conducted in the usual manner either by state operated facilities or by private garages. The diagnostic system will rate vehicles on a five or six point scale and will require action ranging from inspection after another year to immediate removal of the vehicle from the road.

Implementation will take three to five years. The first year will be devoted to detailed system planning, to pre-adoption data collection, and to public education about the program. In the second and subsequent years, the inspection systems will be in operation with the final year emphasizing evaluation and recommendations. A minimum of two years of inspection operation is needed to separate transient and permanent effects. The experiments can occur in three states, preferably ones not currently having a Standard program. More desirably, a single state can adopt the three levels in different areas; the minimum level will be statewide with higher levels located in separate metropolitan areas. This three-in-one approach will be less costly and will minimize for evaluation the effects of differences in population, in environmental characteristics and in administration, but will entail some difficulties in initial planning and in enforcement.

Ultimately vehicle regulation programs seek to reduce the frequency and severity of accidents associated with vehicle components and defects. To accomplish their objective, these programs can manipulate three sets of parameters: vehicle design (usually unavailable to the states), owner maintenance practice, and vehicle inspection. Direct measurement of program impact on accident, injury, and fatality rates has had little undisputed success. An intermediate objective, therefore, is to reduce the frequency of defective components both directly through inspection and indirectly by changing owner maintenance practice. The evaluation procedures recommended principally measure these intermediate effects, with only moderate effort suggested to link the programs to crash reduction.

Quite useful evaluation information can be obtained from the following tools: interviews of drivers during inspections, two types of on-the-road vehicle checks, and analysis of administrative data. Diagnostic sampling of vehicles, at-home surveys of owner maintenance practice, monitoring of automobile replacement parts sales, and analysis of accident trends can provide a more comprehensive evaluation, but only at much higher cost.

Vehicle population sizes, length of experimental period, depth of evaluation, and degree of public cooperation all affect program costs. In a state with three million vehicles, expenditures can range from $6.1 million for a minimum-scope, three-year trial to $14.5 million for a full-range, five-year program. Approximately 40% of these costs will vary directly with the number of vehicles inspected. The degree to which the federal government, the state government, or the motoring public assumes the cost will strongly influence program acceptance.

Local officials will make operational decisions beyond the general outline of the program. Only they are sufficiently familiar with conditions to specify elements such as inspection locations, personnel policies, or detailed operating procedures. Inclusion of these officials in the planning from the start will increase greatly chances for success both by insuring their cooperation and by gaining the benefits of their practical experience.

In the following sections, primary attention will be given to the problems of implementation and evaluation techniques to be used.

5.1 Program Orientation

Programs recommended follow plans that exist in, or that are proposed for, several states. Experimental plans include an on-the-road random check lane in Michigan, Ohio, and California and an owner-conducted self-inspection in Wisconsin (References 5-1 and 5-2). Standard inspection programs already operate in many states, localities, and other nations (References 5-3 to 5-7). Upgrading the standard system to a diagnostic system similar to that used by the Automobile Club of Missouri under NHSTA contract for data collection merits consideration (Reference 5-8). Further, a diagnostic system may be required under exhaust emissions standards. The experimental plan covers these alternatives by instituting a combined check lane and self-inspection program, a standard program and a diagnostic program in areas where a system does not exist.

To date, the inspection systems have not had a complete test of their relative effectiveness. Work has been conducted to measure the effects on mechanical condition of climate, geography, of owner characteristics, of driving conditions, and of the difference between state-operated and state-appointed private garage inspection systems. The results of these studies are not yet available or have been marred by incomplete coverage, biased samples, and non-uniform measurements (References 5-9 and 5-10).

The principal thrust of the present program is to give a controlled test to the various programs. Both "before-and-after" measurements and "among-program" measurements will be attempted. The experiment's objective is to determine which system of inspection will be the most cost/effective in terms of improved vehicle safety quality.

The experiment does not address directly the problems associated with variations within a particular system of inspection. Issues like state-operated vs. state-supervised inspection, like effectiveness of static against dynamic tests of components, and like the development of new techniques for inspecting particular components have been explored in detail or are capable of engineering and laboratory analysis. The three programs will test implicitly some operational differences in inspection technique since each program will require a particular set of methods. The more fundamental issue of which system is in some sense "best" must be resolved, and this resolution can come only by establishing the systems on a controlled basis in actual situations and then observing the results.

The program development process involves five activities: 1) establishing at the Federal level the general outline of the experiment, 2) persuading a state or states to accept the plan, 3) creating operational procedures and cost allocations, 4) conducting operations, and 5) performing evaluation. Each activity will interact with the others in an iterative process. The following discussion covers each activity in turn.

5.2 Establishing the General Experimental Outline

Federal officials must resolve two key issues prior to executing any experiment. They must decide both the number and type of programs to be pursued in particular settings and the duration of program operation. The scientific merit and the cost of the alternatives will determine the direction taken. Program costs are discussed in Section 5.5; ultimately the allocation of these costs among federal, state, and local governments and the public through user charges will depend on negotiations among the concerned government units. Consequently, cost considerations play a secondary role in the discussion of the scientific issues which follow.

The experiments can alter existing standard inspection programs in several states or they can create new programs in different states not currently having a standard or any inspection program. A third means is to develop all three approaches in one or more non-inspection states; this three-in-one approach is recommended as the basic program strategy.

Using states currently having a standard program will produce the problems of amending an ongoing system with a certain built-in inertia and of collecting accurate, current base-line data to determine changes induced by the program. Differences among systems may be smaller than the difference between no system and any system. Thus, these differences will be much more difficult to measure with good statistical precision. Further, allowing some states to change a standard program to the self-inspection and check lane system may contradict previous Federal policy and may generate problems with other states that might want to alter permanently their efforts apart from experimentation.

Placing the programs in three different states which do not currently have standard programs also will generate problems. The problems of securing acceptance will have to be repeated three or more times. The wide variations among states in climate, geography, and population characteristics, which are related to vehicle condition, will tend to contaminate the data needed for evaluation of relative effectiveness. Some variation will also exist among areas in a single state, but this variation is likely to be much smaller than among states in different regions. The three-in-one approach also allows for replications of the basic experiment in several states. Selection of particular areas for observation and experimentation will allow a much closer matching of characteristics than is possible on a state-wide basis. The cost of implementing standard and diagnostic inspections for an entire state will be higher than confining these programs to small areas. The cost savings may be large enough to finance several replications of the three-in-one approach.

Under the three-in-one concept, the state-wide program will be the self-inspection and check lane system. In two counties of the state the standard program and the diagnostic program will be instituted. The selected counties will be in the 100,000 to 400,000 population category and will be widely separated. Much smaller metropolitan areas need to be avoided since they may not contain a representative population mix. Areas of greater than 400,000 population may prove too large to insure uniformity of procedures. Cost considerations in such a large area are likely to become overwhelming in determining the acceptability of the program.

Implementation of the three-in-one approach will entail some administrative problems. First, local officials in the counties selected will have to be persuaded to accept the standard and diagnostic systems. Measures will have to be instituted to insure that motorists in the standard and diagnostic areas do not escape that system for the less demanding system in surrounding counties. Distinctive inspection decals for the standard and diagnostic systems will have to be adopted. Conversely, self-inspection materials will be unavailable to residents of these areas. Police patrols and check lane teams in or near the areas will have to be vigilant to evasion; appropriate penalties will have to be established for evasion.

The time from agreement among federal, state, and local officials to institute the program, to completion of the evaluation, will span several years. How long this period will be is subject both to operational considerations and to desired experimental refinement.

Operational considerations will dictate the minimum amount of time needed to conduct the experiment. It does not seem possible that the program can be accomplished in less than three years. The first year will be occupied with establishing the system. Drafting and enacting legislation and administrative regulations, collecting base-line data for evaluation, recruiting and training personnel, establishing inspection facilities, and informing the public will all have to be achieved before operations begin.

Once operations have been started, another year will be required to completely inspect the entire vehicle population, even in the self-inspection and check lane system. In theory, the self-inspection can be required of everyone by a specific date, but this would overload automobile repair facilities as all vehicle owners sought to meet a single deadline. Check lane teams need to operate for some period of time before they will have affected enough individuals to acquire visibility. In the diagnostic and standard systems, the simple physical limitations of the facilities will dictate an extended period in which to phase-in operations. Thus, the second year will be occupied.

After having all the cars inspected at least once, the experiment could be terminated, but only at the cost of having very incomplete results. In early months, cars which have been inspected under any of the systems can only be located with difficulty. Later when the inspected cars become more common, more data can be collected, but this will then require a crash effort which is not conducive to accurate measurement. Further, observing the condition over a short time period may introduce unexpected seasonal biases in data. Finally, the transition period may not reflect the long-run adjustment of people to the systems. Two effects come to mind immediately; many people may comply more fully with the new regulations simply because of their newness, but later these people may become less active as the system becomes more familiar. Conversely, many inspections may be conducted improperly at first in any of the systems because of inexperience. Examining only the transition period will not reveal these factors, and consequently a third year is indicated to allow normal patterns to develop.

A fourth and fifth year are optional. The additional time will yield the opportunity to collect additional data and will provide assurance that both long-term and short-term effects of the program have been observed.

5.3 Winning Acceptance

The experimental programs will have to be sold to both officials and the public in the areas in which they are conducted. They will demand the cooperation from many groups over an extended period of time. No other countermeasure program directly affects as many individuals on a continuing basis as vehicle regulation. Driver preparation and driver regulation affect the driver infrequently and often indirectly unless he displays a marked pattern of undesirable driving behavior. In contrast, vehicle regulation requires a conscious effort by the owner at annual or semi-annual intervals; this effort often involves his time, his effort, and his money. Consequently, efforts must be made to win acceptance.

In a jurisdiction not having a program, the legislature must enact enabling laws and the executive department must establish the necessary administrative apparatus. Concerns with cost and administrative problems previously has blocked adoption of a vehicle regulation program in several states (References 5-11 and 5-12). Any attempt to implement an experimental program requires a fair amount of preparatory work, and the program must be structured to offer both the promise of permanence and some public benefit. Otherwise, elected officials will be unwilling to antagonize large numbers of their constituents either by adopting a standard program under pressure or by instituting a short-term trial of an alternative program.

Solving the problems of adoption will require effort in four related areas: persuasion of key officials, education of the public, amelioration of opposition, and preparation of detailed procedures. Completion of necessary steps in each area is needed on the part of the sponsoring agency to get a jurisdiction to accept the experimental program.

The first efforts will be directed toward the state officials concerned with highway safety programs. The governor, the governor's highway safety representative, appropriate members of the legislature, the motor vehicle department, and law enforcement agencies must be convinced to include the program in their objectives. Initial contact will be made through the governor's representatives in a number of states. A brief outline of the program will be presented at that time and questions answered. The inquiry will suggest discussion to explore the possibilities of establishing the experimental program in the area. If the state's response appears to be favorable, a meeting involving the groups will be arranged. This meeting will be attended by both regional officials and the personnel who will have principal decision authority over the program. At the meeting the experimental program will be outlined in detail and comments and reactions received. If reaction is again favorable, further consultation concerning specifics can be held with the governor's representatives and others most directly concerned. After review of the information developed from these consultations, a program may be negotiated directly or a request for proposal may be issued to the interested states. The states' proposals will then be reviewed and the necessary awards and contracts negotiated.

In contacting and persuading state officials, several positive points need to be stressed: a firm commitment of federal support for several years; the opportunity for the state to develop the best program in the nation along with the freedom to choose the program which best suits its needs; the improved information that the motoring public will receive; and the sense of security of knowing that vehicles are in better mechanical condition. Negative approaches such as threatening sanctions if no program is adopted, placing too much emphasis on removing junkers from the roads, or claims that failure to act will cost additional lives and crashes, must be avoided.

At some point in discussions with state officials, enlistment of support from non-government groups interested in highway safety will be useful and quite necessary later in winning public support and in providing public education. These non-government groups may well have a degree of rapport and informal contact with state officials which federal officials might not possess.

Once the principal officials at the state level have accepted the program, the general public must be educated about the program. This process has two phases. The first is to persuade them of the desirability of the program. This will primarily be the responsibility of the local officials, who can use the same arguments outlined above. The usual means of public information including mass media, public appearances, and personal contact can be employed. Non-government groups will play a sizeable role here. After the actual program has been adopted, the public must be educated on the specific requirements of the program.

Objections to the program are likely to come from four sources. First to object will be those groups in society which view with distrust any innovative government action. Very little can be done to answer this opposition which is often quite vocal if limited in wide support. Officials can only stress the fact that such programs have been long-established on a sound basis in other areas.

The second source of opposition will come from those who fear that the program will produce graft, corruption, and exploitation of the motorist. These fears unfortunately are somewhat realistic and a number of steps should be taken to ameliorate them. These steps include selection of a state or states that have a reputation for competent administration, and establishment of tight controls which can be employed to avoid these problems. If private agencies conduct inspections they must be carefully policed, and their compensation for inspection must be made great enough to minimize the incentives for improper conduct. State-operated facilities avoid many of these problems, but they do require additional government investment.

A third source of criticism can come from those who feel that inspection will discriminate against poor people who drive the oldest cars (which are usually in the worst condition). Persons who regularly rely on inexpensive used vehicles for basic transportation may not view car repairs as worth the cost, and yet will be unable to obtain well maintained but more expensive vehicles. Two possible benefits of inspection counter this concern. First, the disadvantaged citizens in society are quite susceptable to consumer abuses in general as well as in automobile repairs (Reference 5-13). A fairly conducted inspection program may offer them some additional protection against such abuses particularly if public facilities are made available for inspection on request and means are established to make garages responsible for the repair work performed. However, this could generate opposition from garage operators who normally are quite favorable to inspection programs. As a second benefit, the inspection program may detect minor defects which if corrected promptly may at slight cost prevent much more costly repairs later or greatly extend the useful life of the vehicle.

A final source of opposition may come from those who think that inspection does not offer a return in safety equal to its costs. Potentially this is the most destructive criticism of all particularly in a period when all government programs are coming under question and review. There are strong sentiments to establish priorities for those programs which have a clearly demonstrable effect on the public welfare. Unfortunately, the case for inspection has not been conclusively proven despite many attempts to do so, and presently it may still be impossible to measure the effects of inspection on accidents. The counter argument to this criticism is that the present program is experimental and designed to aid in answering the specific question they are posing. Appointment of an independent agency to perform the evaluation can further blunt this criticism. In this context, establishment of a program advisor and review panel which includes some of the program's critics will not only introduce a valuable diversity of views, but will also reassure critics of the program's integrity.

More space has been devoted to the possible opposition to the program than to the proponents. Those who favor inspection need only be reminded of their own virtue. Establishing a program in an area which does not currently have inspection will, in all likelihood, generate local opposition. The existence and strength of this opposition can be inferred from the absence of a program despite federal standards requiring one. These forces are not likely to be overcome simply by labeling the program an experiment.

5.4 Operational Procedures

In the self-inspection system, vehicle owners will receive instruction manuals and certification forms at some pre-determined time during the initial year of operations. The instruction manual will contain a complete guide on how to conduct the self-inspection along with suggestions for a more complete inspection than the minimum if the owner wants to do one. The owner or, if he is unwilling or is unable, a mechanic of his choice, will complete the required inspection indicating, on the certification form provided, the defects found. The owner will then repair the defects, provide a copy of the form indicating the defects found and certifying repair to the agency supervising self-inspection. The owner will also retain one copy of the certification form in the vehicle and place a decal indicating when the vehicle is required to be reinspected on the vehicle.

Distribution and return of certification forms will be dependent on the motor vehicle registration system used in the state. If all registrations are renewed during one period, the manuals will have to be distributed by mail during the year and the certification forms returned also by mail. If the owner does not return the certification form within the required period, a follow-up procedure will be invoked. If the state follows a continuous registration procedure, the distribution and follow-up will be simpler. The owner will receive the necessary materials shortly prior to the expiration of his registration. He will then be required to complete the inspection and present necessary certification before receiving his new license plates. Transfer of vehicles from one owner to another or purchase of new vehicles will require the new owner to complete and certify inspection prior to receiving the title and registration. In all cases, the owner will receive a manual and sufficient forms to complete inspection for several years.

Unfortunately, simply requiring vehicle owners to inspect and to certify their vehicles may not be enough to insure a high quality level of mechanical condition. A police operated check lane system will complement the self-inspection program by enforcing its provisions and by requiring a continuous adherence to a minimum quality level.

Check lane teams consisting of two police officers and two vehicle inspectors will be operated at various sites throughout the state. Vehicles will be selected from the traffic stream for inspection. The inspection will determine if the self-inspection had been conducted and where possible if the repairs had in fact been made. The vehicle will also be checked for defects which could have occurred after the self-inspection was completed. If any defects are found, the self-inspection sticker will be removed from the vehicle and a limited-time failure sticker attached. The owner will then be required to have the vehicle repaired, and he must return to a designated place, such as a state police post, to have his vehicle reinspected. If the defects are corrected or if none were present at the original inspection, a new check lane sticker will be attached to the vehicle. This sticker will be valid until the time that the self-inspection certification will have expired. Cars displaying check lane stickers will not have immunity from further check lane inspections, particularly if they display obvious hazardous defects, but presumably the police will concentrate their efforts on those not previously contacted. In cases where no self-inspection was performed, where the self-inspection was performed but the repairs were not made, where hazardous defects had occurred after self-inspection repair,,or where the owner failed to return for reinspection, criminal penalties may be invoked.

Vehicles may be checked also in a less systematic manner. Officers making routine traffic stops may be empowered to conduct inspections similar to that given in the check lane with the same procedures being followed if any irregularities are discovered. When vehicles with hazardous defects are observed by police, they will be required to institute the defective vehicle procedure of the check lane.

A reasonable goal for the check lane program will be to check approximately 10% of the vehicles per year on a state-wide basis. The intensity of the effort can be varied depending on judgment and experience as to the relative degree of compliance and condition of the vehicles in a particular area. Areas which had poor compliance records or had vehicles with a high frequency of multiple defects will be checked more frequently, while areas with good compliance will be checked less frequently but often enough to remind citizens in that area that everyone is subject to the law. Vehicles displaying new certification stickers will be given particular attention since improper certification is most easily detectable in these cases. From time to time publicity should be given to incidents in which local residents encountered difficulty with the check lane to serve as a reminder to others.

Having the check lane will encourage individuals to maintain their vehicles between inspection periods. Along with the educational value of self-inspection, the check lane can induce a continuous awareness of vehicle condition which can alter maintenance practice in such a manner as to have a more beneficial effect on the quality of mechanical condition than an inspection periodically conducted by someone else. The check lane will also have useful side benefits in terms of greater enforcement of driver's license restrictions or suspensions, of recovery of stolen vehicles, and of general law enforcement visibility (Reference 5-14).

Little needs to be said about the organization and administration of the standard inspection program. There are a variety of systems in existence, many of which provide adequate models for an experimental jurisdiction to follow. For the experiment, adoption of a publicly operated facility seems more desirable. This will insure that a more consistent procedure will be followed and will ease the burdens of collecting data for evaluation. Moreover, by having a publicly operated facility, fewer special interests will develop favoring its continuation if some other system is found preferable after evaluation. A public facility will entail greater government outlay and will place some additional costs on the public; many people will have their vehicles checked first at a private garage before going to the public station so as to avoid having to return for a reinspection.

The diagnostic inspection system is more complicated than either the self-inspection system or the standard inspection system; it requires a more structured inspection apparatus and involves multi-level set responses which will be more difficult to administer. In the diagnostic system more individual components will be checked and their performance measured than in other systems. Ratings will be given to components individually, and to the vehicle as a whole. A critical assessment will be made of the deviation of each component from specifications. A five point scale is recommended with varying actions required of the motorist. The suggested scale and recommended responses are:

A. For component, excellent condition, shows little or no wear, within original equipment specifications. For vehicle, all systems are "A" rated. Reinspection required after one year.

B. For component, normal condition or normal wear, within good tolerance of original specifications. For vehicle, all systems are rated "B" or better. Reinspection required after one year.

C. For component, excessive wear shown, but does not represent a hazard or potential hazard under extreme conditions. For vehicle, all systems are rated "C" or better. Reinspection required in six months, but repair only advised not required.

D. For component, excessive wear shown or failure that will constitute a hazard under extreme operating conditions or potential hazard under normal operating conditions. For vehicle, any system with a "D" rating or numerous systems with "C" ratings so as to indicate generally poor maintenance practice. Reinspection required in one month. Repair required.

E. For component, excessive wear shown or failure that will constitute a hazard under normal operating conditions. For vehicle, any system with an "E" rating or with multiple "D" ratings so as to indicate generally poor maintenance practice. Reinspection required in two weeks. Repair required. Vehicle operation restricted to certain circumstances as from inspection facility to garage and to and from work if no other means of transportation available.

Condemned. For component, excessive wear shown or failure. Constitutes an immediate threat to safety. For vehicle any condemned system or multiple "E" rated systems which in combination pose immediate threat. Repair required. Must be reinspected before any highway operation is allowed.

What constitutes a particular rating level for each component will have to be decided in the planning phase of the specific diagnostic facility.

The diagnostic system alters some of the risk factors associated with inspection systems. It will reduce the chance of some critical component being overlooked by a more thorough and detailed examination. But it will allow certain components to "pass" by the mechanism of the "B" and "C" ratings. These components might be rejected under a strict go-no-go system of extremely cautious tolerances. By advising the owner and by requiring reinspection sooner in the case of "C" ratings, the risk of this action is reduced while at the same time eliminating the need for repairs of marginal benefit.

The varying levels of response will allow the vehicle owner to form an accurate assessment of the condition of his vehicle and to tailor his maintenance policy to that assessment. The owner will have to repair only those components that need repair rather than perhaps being forced to replace an entire sub-system. He will be alerted to excessive wear of particular components that may cause more extensive damage if not corrected. Finally, he will dismiss the notion that having passed inspection means a perfect car that obviates the need for continued caution.

5.5 Program Costs

Program costs will vary with the number of vehicles to be inspected; the extent of inspection; the degree of public acceptance of the program in fulfilling requirements without enforcement action being taken; the type and classification of personnel used in each program activity; local wage levels; and the productivity of personnel. Detailed cost estimates can only be constructed after a final program plan has been decided for a specific jurisdiction. Even these estimates should be prepared for a range of anticipated operating results: for example, costs can vary widely with number of reinspections necessary in the standard or diagnostic program and with the operating efficiency of the lane.

To illustrate some of the problems associated with determining program costs, estimates have been prepared for an experimental program under several "reasonable" assumptions. They have not been drawn on the basis of detailed examination of any one system, but rather upon judgment based on observation of a number of systems. The cost estimates presented represent perhaps the 25th percentile and the 75th percentile of costs. Put another way, the estimates represent what might prevail under moderately favorable or under moderately unfavorable conditions. For evaluation, the estimates also reflect the differences between a modest and a large effort in data collection. The calculations assume a state with 3,000,000 registered vehicles using a self-inspection and check lane system operating state-wide and having standard and diagnostic systems operating in two counties with approximately 100,000 vehicles in each. More detailed assumptions for each program category are presented below. Table 5-1 presents these estimates for both a three and five year program.

The Federal portion of the cost of each activity assumes that the Federal government will finance 100% of the initial and capital costs and 70% of the operating costs of each program. The Federal government is assumed to be funding 100% of the evaluation cost. This mix of funding is purely arbitrary; it might arise from a combination of 403 and 402 funds with some matching by the state.

No attempt has been made to estimate directly the vehicle owner's compliance cost, a major cost element of vehicle inspection programs which does not appear in governmental budgets. This compliance cost includes the value of the owner's time, the incremental repair cost to meet inspection standards, and miscellaneous items such as vehicle operating expense while traveling to and from the inspection facilities. Since these expenses are subject to both subjective valuations and a large stochastic element no attempt has been made to formally estimate them; a purely subjective estimate places this cost between $2.50 and $7.50 per inspection.

The low range estimate for the check lane system, presented in Table 5-2, was based on the assumption that there would be 18 teams operating, each inspecting 100 vehicles per day for 175 operating days per year. Each check lane team will consist of two police officers, and two civilian vehicle inspectors. A third vehicle inspector will be assigned full time to performing inspections at a central location. On free days, the two check lane

Table 5-1A

Vehicle Regulation Program: Minimum and Maximum Costs by Major Activity

Activity	Total Costs Minimum	Total Costs Maximum	Federal Costs[1] Minimum	Federal Costs[1] Maximum
Check Lane	$2,413,000	$5,005,000	$1,772,000	$3,323,000
Self-Inspection	2,097,000	4,339,000	849,000	3,146,000
Standard Inspection	314,000	748,000	80,000	105,000
Diagnostic Inspection	1,072,000	3,158,000	833,000	2,515,000
Evaluation	150,000	895,000	150,000	895,000
Totals	$6,046,000	$14,145,000	$3,684,000	$9,984,000

Table 5-1B

Vehicle Regulation Programs: Total Costs of Three and Five Year Programs for Both Minimum and Maximum Program by Major Activity

Activity	3 Year Program Minimum	3 Year Program Maximum	5 Year Program Minimum	5 Year Program Maximum
Check Lane	$2,389,000	$2,639,000	$4,550,000	$5,053,000
Self-Inspection	2,097,000	2,839,000	3,089,000	4,339,000
Standard Inspection	314,000	535,000	528,000	748,000
Diagnostic Inspection	1,072,000	1,844,000	1,804,000	3,158,000
Evaluation	150,000	530,000	256,000	895,000
Totals	$6,022,000	$8,387,000	$10,227,000	$14,293,000

Note: Numbers presented here might not agree exactly with those in Tables I-VI due to rounding and certain minor adjustments for joint overhead costs shared among certain activities.

[1] Federal expenditures are assumed to cover 100% of capital and set up costs for all programs, 70% of operating costs not covered by user fees for inspection programs, and 100% of evaluation costs.

Table 5-2

Vehicle Check Lane, Minimum and Maximum Program Costs for Three and Five Year Programs

Minimum Program[1]

Item	3 Year Cost[2]	5 Year Cost[3]
Program Planning	50,000	50,000
Personnel Training	150,000	150,000
Equipment	108,000	198,000
Personnel:		
45 Officers at $10,000/yr.	900,000	1,800,000
54 Inspectors at $7,000/yr.	810,000	1,620,000
4 Supervisors at $12,500/yr.	100,000	200,000
4 Clerks at $6,000/yr.	48,000	96,000
Transportation and Vehicle Upkeep	43,200	86,400
Data Processing and Supplies	126,000	252,000
Office Space and Other Overhead	78,000	156,000
Total Operation	$2,437,200	$4,646,400
Less Adjustment for Evaluation Activity:		
150 Team days per year of operations at $60.00/day	18,000	36,000
Data Processing and Supplies	6,000	12,000
Total Chargeable to Inspection Activity Only	$2,413,200	$4,598,400
Federal Share		
Start-Up and Capital	308,000	398,000
70% of Operations Excluding Evaluation	1,474,000	2,940,000
	$1,772,000	$3,338,000

[1]Minimum program consists of 18 teams operating 175 days per year in field and inspecting 100 vehicles per day. 150 team days per year are devoted to collecting data for evaluation. Evaluation program is charged for the variable costs of this activity which are $60.00 per team day for personnel and $ .20 per vehicle inspected for data processing.

[2]Three year program involves one year of program planning and organization and two years of operation.

[3]Five year program has one year of planning and four years of operations.

Table 5-2 (Continued)

Maximum Program[4]

Item	3 Year Cost	5 Year Cost
Program Planning	$ 50,000	$ 50,000
Personnel Training	165,000	165,000
Equipment	110,000	210,000
Personnel:		
50 Officers at $10,000	1,000,000	2,000,000
60 Inspectors at $7,500	900,000	900,000
4 Supervisors at $12,500	100,000	100,000
4 Clerks at $6,000	48,000	96,000
Vehicle Upkeep and Transportation	48,000	96,000
Data Processing and Supplies	140,000	280,000
Office Space and Overhead	78,000	156,000
Total Operations	$2,639,000	$5,053,000
Less Evaluation Allocation:		
Inspection Personnel	200,000	400,000
Data Processing	20,000	40,000
Inspection Operations	$2,419,000	$4,613,000
Federal Share		
Start-up and Capital	325,000	425,000
70% of Operations	1,465,000	2,932,000
	$1,791,000	$3,357,000

[4]Maximum program involves 20 teams operating 175 days per year with 24 days per team devoted to evaluation activities. Each team will conduct 25 evaluation samples per year. Inspection personnel charges are higher proportionally than under minimum plan for evaluation effort since the evaluation effort requires a substantial rather than a minor effort.

Table 5-3

Self-Inspection Program Costs by Major Program Category, Minimum and Maximum Estimates for Three and Five Year Programs

	Minimum Program[1]		Maximum Program[2]	
Category	3 Years	5 Years	3 Years	5 Years
Initial Planning and Design	$ 30,000	$ 30,000	$ 30,000	$ 30,000
Distribution of Original Materials to Public	750,000	750,000	1,050,000	1,050,000
Renewal and Replacement of Materials	300,000	600,000	410,000	820,000
Processing Returned Certificates	900,000	1,500,000	900,000	1,500,000
Follow-Up of Deliquents	117,000	209,000	449,000	939,000
Total Cost	$2,097,000	$3,089,000	$2,839,000	$4,339,000

[1]Minimum program assumes that transactions can be handled through regular registration procedures and that certification must be presented to receive new license plates.

[2]Maximum program assumes that transactions are handled independently of regular registration.

Table 5-4

Standard Inspection Program Costs, Minimum and Maximum Estimates for Three and Five Year Programs

	Minimum[1]		Maximum[2]	
Category	3 Years	5 Years	3 Years	5 Years
Planning and Design	30,000	30,000	30,000	30,000
Construction and Equipment	50,000	50,000	75,000	75,000
Operating Personnel	174,000	348,000	254,000	538,000
Supplies and Maintenance	60,000	100,000	81,000	135,000
Total Operations	$314,000	$528,000	$535,000	$748,000
User Charges	$223,000	$448,000	$430,000	$643,000
Government Cost	$ 80,000	$ 80,000	$105,000	$105,000
Federal Cost				
Start-Up	$ 30,000	$ 30,000	$ 30,000	$ 30,000
Capital	$ 50,000	$ 50,000	$ 75,000	$ 75,000

[1]Minimum program assumes that two facilities will be required.

[2]Maximum program assumes that three facilities will be required.

Table 5-5

Estimated Costs for Diagnostic Inspection Program by Principal Cost Category, Minimum and Maximum Estimates, Three and Five Years

	Minimum[1]		Maximum[2]	
Category	3 Years	5 Years	3 Years	5 Years
Planning and Design	$ 50,000	$ 50,000	$ 50,000	$ 50,000
Construction and Equipment	280,000	280,000	480,000	480,000
Operating Personnel	576,000	1,152,000	1,844,000	3,158,000
Records and Data Processing	66,000	131,000	75,000	150,000
Materials and Supplies	105,000	201,000	180,000	360,000
Total Operation	$1,072,000	$1,804,000	$1,844,000	$3,158,000
Federal Cost				
Capital and Start-Up	330,000	330,000	530,000	530,000
Operations	503,000	1,026,000	884,000	1,985,000
Total	$833,000	$1,356,000	$1,414,000	$2,515,000

[1] Minimum operation involves use of 8 facilities costing $35,000 to construct and processing 12,500 vehicles per year.

[2] Maximum operation involves use of 12 facilities costing $40,000 to construct and processing 10,000 vehicles per year.

Table 5-6

Estimated Costs for Evaluation Program by Major Program Elements. Minimum and Maximum Estimates, Three and Five Years

	Minimum		Maximum	
Category	3 Years	5 Years	3 Years	5 Years
Professional Personnel	98,000	138,000	122,000	182,000
Data Collection:				
Lighting Survey	14,000	23,000	14,000	23,000
Driver Interview	7,500	12,000	7,500	12,000
Check Lane Sample	24,000	48,000	220,000	440,000
Accident Data	15,000	25,000	15,000	25,000
At-Home Survey	----	----	60,000	80,000
Diagnostic Sample	----	----	68,500	95,500
Business Survey	----	----	22,500	37,500
Total Operation	$150,500	$256,000	$529,500	$895,000

inspectors will assist with reinspections. The high-range estimate for the system assumes that 20 teams will be required to operate in the field only 150 days per year. The remainder of their time will be devoted to reinspection activities and to data gathering for evaluation. This latter activity is charged, though, to the costs of the evaluation. The three year and five year costs include initial start-up costs, capital costs, and expenditures for two and four years of full operations. No allowance is made for the costs of personnel turnover, nor do the salary estimates include much allowance for fringe benefits.

Costs of the self-inspection programs have been estimated in a fairly arbitrary manner. The operation is viewed as consisting of a number of transactions between the motor vehicle owner and the agency operating the program. Unit costs are assigned to each transaction, and the number of transactions necessary are estimated. These estimates are presented in Table 5-4. The lower costs will prevail if the plan ties inspection to a continuous renewal registration procedure. Under this arrangement principal cost elements will be preparation and distribution of materials in the initial year, processing of returned certification forms, and handling of approximately 600,000 transfers of ownership annually. Higher costs will be obtained if the plan is largely independent of the registration procedure. In addition to the costs stated above, follow-up procedures must be taken for those owners who fail to return certification forms within the allotted time. Use of two follow-up letters and referral to the check lane team for contact is anticipated. The need for follow-up can grow over time as people become more accustomed to the program and less careful in meeting the requirements. Estimates were based on the following pattern: first year, 90% no action required, 5% one letter, 4% second letter, 1% check lane contact; second year, 70% no action required, 20% one letter, 8% second letter, and 2% check lane contact; third and subsequent years, 50% no action required, 30%

one letter, 15% second letter, and 5% check lane contact. These efforts for follow-up and complete certification are necessary to insure a means of monitoring the program and to impress the public with their responsibilities. Allowance is made for the expense of preparing materials and establishing administrative procedures in the initial year.

Data for the costs of the standard inspection program are shown in Table 5-5. These estimates are for a governmentally operated facility in a community with approximately 100,000 vehicles. The low-range figures assume that each unit can process 250 cars per day and that 20% of the cars will be required to return for re-inspection after repair. The high-range estimate assumes that each unit can inspect 200 cars per day and that 30% of the vehicles will have to return for reinspection. Capital costs are assumed to run at $20,000 per unit for the low-range estimate and at $30,000 for the high-range estimate. Little or no allowance is made in either the three or five year program estimates for replacement of inspection equipment, retraining of personnel, or for many clerical and support expenses. Federal expenditures are limited to the capital costs since inspection fees of $1.50 to $1.75 will cover the direct operating costs of the facility.

Costing of the diagnostic system is presented in Table 5-5. These costs are influenced by the detail of diagnosis performed and also on such elements as the productivity of the inspectors, the frequency of follow-up inspections, and the detail and extent of record keeping. In the low-range estimate, the lanes were assumed capable of inspecting 125 vehicles per day with 20% of the vehicles requiring reinspection within six months, and 10% of the vehicles requiring immediate reinspection after repairs were made. Capital costs were assumed to be $30,000 per lane. In the high-range estimate, the lanes were assumed capable of inspecting 75 vehicles per day with 30% of the vehicles requiring reinspection within six months and with 15% of the vehicles requiring immediate reinspection after repairs were made. Capital costs were assumed to be $40,000 per lane. Again, the base was assumed to be 100,000 vehicles. Inspection fees will be established at the same level as in the standard system. The difference between the fee income and the operating costs will be shared on the 70%-30% federal-state arrangement assumed for other activities.

Discussion of evaluation costs is incomplete without reference to the next section. The data in Table 5-6 represent two ranges of evaluation. The low-range assumes a limited effort using a few measures which are relatively easy to administer with use of untrained personnel. A rather limited analysis effort will be conducted by both senior and junior researchers working on a part-time basis. The high-range estimate assumes use of a wider range of instruments administered in many cases by trained personnel. The analysis effort has more full-time effort by both senior and junior researchers with a greater complement of individuals at both levels working during critical phases of the program.

5.6 Evaluation

Evaluation can be performed with a number of procedures which yield both direct and indirect measures of changes in vehicle condition, owner maintenance practice, and accident involvement. Changes in vehicle condition and in owner maintenance practice can be measured with good precision, but differences in accident involvement resulting from the programs will be quite difficult to detect. A subsidiary goal will be to examine public attitudes toward the inspection program.

Four direct measurements of changes in vehicle condition can be used: inspections conducted by the check lane teams under controlled conditions, visual observation of vehicles on the roads, a diagnostic examination of a random sample of vehicles, and statistical analysis of administrative data from program operations. At-home surveys of vehicle owners and interviews of persons during inspection can provide both direct information on, and indirect inferences of, changes in maintenance practice and acceptance of the program. Collection of data on automobile parts sales may provide an indirect measure both of inspection effects and of changes in owner repair practice. Statistical analysis of police accident reports in conjunction with information on the inspection status of the involved vehicles may offer a measure of accident involvement effects.

In outlining an evaluation plan, several questions have to be considered. How many procedures will be used? On what base populations will the procedures be applied? How many factors will be included in detail? What statistical precision will be obtained by the estimates? The exact program followed will represent a compromise between depth, accuracy, and cost of the experiment. A range of alternatives is suggested: a small number of procedures will provide enough data to insure that the experimental effort will not be wasted; use of all the procedures will provide much more precision but approach the range where diminishing returns will occur. The smaller effort will concentrate on the two counties with standard and diagnostic systems plus one additional county with the self-inspection and check lane system. The larger effort will attempt to measure in some detail changes in the entire state. Both levels of effort will concentrate on estimating the main effects of the programs on the population as a whole, rather than placing specific attention on differential effects among various population sub-groups.

In the ideal experiment, one would attempt to measure the effects of the programs on a large number of interdependent measurements of the vehicle's condition. Further, the experiment would be closely controlled for a number of factors which influence vehicle condition. These elements include the distribution of vehicle age, vehicle mileage under different operating conditions, several owner characteristics, and miscellaneous climatalogical and environmental factors. To perform a complete factorial experiment would involve thousands of observations simply to cover all possible combinations of effects. A compromise between this degree of statistical refinement and cost has to be accepted. What is proposed is to use a number of measures which will have large enough sample sizes to guarantee that large changes such as a reduction of outages of a particular component from say 20% to 15% will be detected with fairly high confidence. The matching of experimental areas on population and environmental characteristics and the use of vehicle age, mileage, and type classifications in several of the measurements will reduce the possibilities of bias due to incomplete specification. Still the precision achieved by the suggested program will be less than that which can be achieved by a more ambitious one. Yet, one may argue that unless the effects produced by the experiments are substantial enough to be captured by the present measurements, the effects will have little influence on the chain of accident causation.

The first evaluation technique is a controlled use of the check lane teams to gather data. Under this plan, each team will visit four pre-selected sites on six occasions each year and will conduct inspections of vehicles which will be randomly selected from the traffic stream. The four sites will be categorized by income and population density of the area surrounding them. The six visits per site per year will yield a total annual sample of 2400 vehicles assuming 100 vehicles inspected per site visit. In the full program, all of the check lane teams will participate in the data gathering effort. In the smaller program, only teams covering the three study counties will gather data. This methodology assumes that vehicles in traffic at any particular point in an area will be representative of the general vehicle population in that area. Careful analysis must be made of the age, make, and type distributions of the vehicles sampled to insure that this assumption is statistically valid.

Site selection for this particular measurement must be done with some care. The area from which the samples are to be taken must be divided into sub-areas either on the basis of census tracts or arbitrary geographical units such as townships. These sub-areas will be cross classified by income and population density and an area will be selected from each category at random. Once the particular sub-areas have been designated, they will be surveyed and several possible inspection sites will be selected in each sub-area. The inspection sites will have high densities of local traffic and will have the necessary physical features to accommodate safely the team's operation. At least 10 such sites should be designated; one of those sites will be selected by chance to be the sampling location. Some judgment must be applied in drawing the list of candidate locations so as to insure that the points chosen do not yield data from an excessively small area lest the repeated presence of the check lane team in its sampling capacity overly influence the behavior of area residents.

The procedure recommended follows closely the pattern used by survey research organizations for selecting households to interview. A city will be divided into areas by certain population characteristics. A number of these areas will be selected by chance, and then particular addresses and households will again be selected by chance. Rather than picking a resident of an area at random on the basis of address, the check lane method randomly selects residents of an area who happen to be out and about at the time.

Two dangers exist in the technique. First, despite the care used in selecting locations, the particular location may not be representative of the vehicles in the general area due to the influence of factors which are not revealed by analyzing their make, age, and type distributions. This is a risk present in any stratified probability sample. Second, it will be possible for a team anxious to "make a record" to contaminate the experimental results either by regularly working more in the sample area than in other areas or by working intensively in the sample area immediately prior to the time it is to conduct the sample. This can be prevented by not permitting regular operations within a certain radius of the sample site, say five miles, and by not informing the team that it is to take a sample at a particular location until the designated day.

Data gathered from the sample locations would be processed in the same manner as the inspection reports from the regular check lane activity, but this information will be passed to the evaluation group for statistical analysis. Hopefully, the data would show decreases in the outage rates over time as the different systems took effect.

A second means of gathering data about vehicle condition is to observe vehicles in some systematic manner from the roadside without stopping them. In this plan observers are stationed at night either at stop-sign controlled or traffic-signal controlled intersections. They can then record the frequency of defects in the lighting systems, exhaust systems, and possibly the glass of the vehicles which stop for the traffic control device. Under the proposed experimental plan eight sites will be selected in three counties: the two intensive inspection counties, and a control county from the self-inspection program. The sites will be selected in essentially the same manner as in the sample check lane operation. Each site will be visited once a month over the course of the experiment with the expected sample sizes being between 150 to 500 vehicles per site visit.

This technique may be quite useful in detecting changes in some common forms of vehicle defects which may be quite sensitive to inspection procedures. The measure has the advantage of being relatively inexpensive per data observation. It can be conducted in an unobtrusive manner quite independently of any administrative function of the program. Non-police personnel may be used and the cooperation of the public need not be secured. It will complement the operation of the sample check lane program covering vehicles at a time not normally available for check lane operations.

The method is subject to the criticism that there is no way of guaranteeing that the vehicles observed are in any way representative of the vehicle population; no way exists even to check on biases. An earlier, and as yet unpublished, HSRI study indicates that the technique does measure large differences between particular sites. These differences consistently follow income and population density patterns, and data from any particular site or for closely adjacent sites in the same type of area were very consistent over time. This experience indicates that the procedure may be quite accurate despite its obvious simplicity.

The program will provide approximately 2400 vehicles per month per county. With sample sizes this large quite precise estimates may be obtained of small differences in vehicle condition. Data gathered in the technique will be analyzed by the usual statistical methods, and the appropriate comparisons will be made between the self-inspection, standard inspection, and diagnostic inspection areas over time. The form used to collect the data in the previous study is shown in Figure 5-1 and a copy of the instructions for the experiment is included in Appendix A.

Diagnostic sampling of the vehicle population provides a third data source. Five hundred vehicles will be randomly selected from the registration records of each of the three study counties and will be processed through a portable diagnostic facility similar to that developed in a recent NHTSA research program. In a three-year program, the sample will be repeated each year of the program, once before inspection begins, once in the transition year, and once in the first year of complete inspection. For a five-year program an additional sampling will be conducted in the final year to determine any further long-term effects.

Figure 5-1

AUTO LIGHTING SURVEY DATA SHEET

Study No ________
(2-5)

Date ____________ Page ____ of ____

5-8)	Car Number									
-11)	Make									
-13)	Year									
(14)	R head out									
(15)	R head brk									
(16)	L head out									
(17)	L head brk									
(18)	F glass									
(19)	S glass									
(20)	B glass									
(21)	Muffler									
(22)	R brake ot									
(23)	L brake ot									
(24)	Both bk ot									
(25)	R tail ot									
(26)	R tail brk									
(27)	L tail ot									
(28)	L tail brk									
(29)	Both t ot									
(30)	Both t brk									
(31)	Plate out									
(32)	Wiring									
(33)	Other									
(34)	Outstate									

HSRI Form 21

The sample size of 500 cars was determined both by the capacity of the facility and by statistical criteria. Working for a month, the facility can possibly obtain 25 vehicles per day for 20 working days. By limiting operation to a month in a particular location, the three counties can be covered in close enough time proximity to minimize seasonal effects. The sample size of 500 is sufficiently large to insure that relatively small differences in the frequencies of the common defects can be observed with high confidence. In selecting the vehicles, the registration records will be stratified by make and by model year classifications to insure an unbiased sample. With the small sample size, it is unlikely that any significant information will be gained about the differential effects of the programs on the various vehicle sub-groups. Still information will be gathered on these parameters, and the data will be explored to see if any effects are significant.

Once the 500 vehicles have been selected, every effort must be made to inspect those original vehicles. Differences between people who are willing to cooperate and those who do not often represent an important source of bias. An appropriate scheme for insuring complete coverage will be to contact initially all selected owners by mail and request that they participate. Those who do volunteer and do bring their cars in for inspection may constitute the majority. Those who do not volunteer will be contacted by phone, and the program will be explained to them in detail. If they do not respond to this second request, they should be offered a nominal fee, say $10, to have their vehicles inspected. If this reward does not prove sufficient, a larger fee should be offered, say $25. If this does not work, it may be advisable, providing cooperation is obtained, to use the legal power of the inspection system to force the persons to have their vehicles checked. This must be used as a last resort and can lead to legal difficulties not only for the sampling but for the entire program. Obtaining as complete coverage as possible from the sample is critical to this particular activity considering the sample sizes involved and the high cost of the program. To contaminate the sample with a large non-response bias will render the costly results almost meaningless.

The signal advantage of the program is that it can provide a consistent measurement, based on a true random sample, of the differences among the three systems which is not offered by any other available measurement tool. The diagnostic procedures will duplicate as closely as possible those used in the diagnostic inspection system. The disadvantage of this approach is its relatively high cost which leads to a small sample size. If the evaluation must operate under a limited budget, the cost disadvantage will force the abandonment of this measurement.

A fourth source of data on vehicle condition is the administrative information compiled in the course of inspection operations. In evaluation, this data will be checked for any major trends over time among the various inspection formats. In the initial phase of the program, administrative data will provide a source of information on how rapidly the system is being implemented. It may be used to check the validity of other data gathering activities.

However, not too much reliance can be placed on administrative data as an evaluation tool. Owner self-inspection reports will contain inaccuracies and will reflect differing interpretations of the regulations by a large number of persons. Reports from normal check lane operations contain a number of biases. Effective enforcement policy will require that activity be concentrated in problem areas for a sufficient period of time to inspect a large number of defective vehicles. Officers will also become more skilled in detecting defective vehicles. Reports from standard inspection stations and from the diagnostic facilities will show two important biases. First, vehicles which are checked at these facilities will in some sense represent the worst of the population since they will have been longest without inspection. Contrariwise, many vehicles which are presented to the inspection stations may have undergone a pre-inspection so that the owner can complete formal inspection without worrying about returning for a re-inspection. Consequently, all sources of administrative data will be contaminated and should not be given too much weight in evaluation, but since it needs to be collected in the normal course of affairs and is available at minimal additional cost it should not be ignored in conducting the evaluation.

The second major goal of evaluation is to measure changes in owner repair practice induced by the inspection systems. Such changes can either enhance or dilute the program's direct effect of forcing repair of defective components. If owners become more aware of and concerned with the condition of their cars, their increased maintenance acitivities can greatly enhance the vehicle safety quality level; this will occur by causing defects to be detected and repaired more quickly and by extending the useful life of many components through preventative maintenance. Conversely, many individuals may relax their own efforts and rely on the inspection to detect faulty items, or they may postpone repairing faults until the time of inspection.

The most practical way to determine what changes have occurred is to ask people to describe their behavior in some detail. This must be accomplished both before the program is instituted and after operations have been in effect for some time, since reports of current behavior are most likely to be accurate. The questions asked must be both direct and indirect. The direct questions are posed in hopes of obtaining some objective measurement of behavior, but the responses can be clouded by the person's desire to give what he thinks is the socially acceptable answer. Indirect questions are used to counter this problem.

In the suggested study, the interviewing will cover three principal areas. Owners will be questioned about their repair activities in the recent past and asked to assess the condition of their vehicle. The responses will be used as a direct measure of changes in behavior. As an indirect measure, questions will be asked concerning individual's knowledge of, and attitude toward, maintaining their vehicle. The hypothesis is that improved knowledge of vehicle components and a more positive attitude toward maintenance will be reflected in behavior. A subsidiary goal will be to check peoples acceptance of the program in their area under the assumption that a more popular program will enhance chances for that program's success. Responses to these questions will be related to factors which include the person's age, sex, education, driving experience, and specialized knowledge or training in automobile mechanics.

To explore some of the issues raised above a sample questionnaire has been prepared and is presented in Appendix B. This questionnaire need not be used in its entirety and some of the items will have to be altered to fit local conditions. Since the questionnaire has not been validated, a pre-test will be needed early in the program planning. The questionnaire has five sections: a) General information about age, sex, education, and driving experience, b) Owner's present repair practice, experience, and estimate of vehicle condition, c) Attitudes about vehicle maintenance and safety, d) Knowledge of automobile mechanics, and d) Opinions about automobile inspection both in general and in reference to the local system.

Owners and drivers can be contacted in two types of setting, during inspection procedure or at-home independently of inspection. In both settings, a common set of questions will be used as the basis for the interview. Some additional questions can be asked with greater success in a particular setting. Information about attitudes on inspection can be gathered more easily during inspection while the owner is most aware of his feelings. Much more detailed information on other items can be obtained when the driver or owner is at ease in his own home.

In the inspection setting, the owner will complete a questionnaire while his vehicle is being checked. A copy of the inspection report for the vehicle will be attached to each completed or refused questionnaire so that the owner's responses can be related to observed vehicle condition. Questionnaires will be distributed to patrons of the diagnostic and standard systems one day a month. In the check lane program, interviews will be conducted on those days when the team is operating in the sampling mode. Using these sampling frequencies, 2400 observations will be made per system per year. These frequencies will be statistically adequate to determine relatively small differences in response. Interviewing persons during inspection presents two difficulties. The drivers of the vehicles inspected may not be the same persons who are normally concerned with maintenance. Secondly, individuals may not be candid in their responses if they feel that somehow their responses will affect the outcome of the inspection. This may be particularly true in the check lane.

Interviewing vehicle owners apart from inspection is the second setting for gathering information. For both a three-year and a five-year experiment, a minimum of two surveys will be conducted. One survey will be before inspection operations begin, and the second survey will occur in the last year. One or two additional surveys can be conducted if funding permits. The most useful time for the third survey will be in the transition year while people are adjusting to the new system. In each area, 800 respondents will be identified by random sample of vehicle registration records; this will give a total of 2400 respondents for each survey which yields high statistical precision.

The most productive method to conduct the survey is to interview the person in his home. The person is contacted initially by letter and arrangements are made to visit him at a convenient time. If the person does not initially agree, further contacts are made until he does agree or it becomes apparent that he will not cooperate despite the most persistent efforts. By using the at-home interview, the survey can ask the most detailed questions and will usually receive the most thoughtful and careful responses. Further, the interviewer can establish rapport with the subject which will lead to a more successful completion of the survey; in this, the interviewer can form judgments about the candor and interest of the respondent which will greatly aid in analyzing the data. The principal disadvantage of this approach is its cost. Interviewing the subject by telephone is less costly, but at a minimum these cost savings will be at the expense of obtaining detailed information. Few people are willing to answer extensive questions on the phone particularly when they have no means of verifying the identity of the questioner. The interviewer will have no way of forming the judgments about the respondent that a face to face meeting can produce. The telephone method should be used only if funds are not available for any other form of contact. A final technique, which would save some cost, is to tie the driver interview with the diagnostic sampling of the vehicles. This has some promise and can be tested in the early part of the program. The principal limitation of the method is that it will make the interview subject to the same difficulties of obtaining cooperation as are anticipated for the diagnostic sample. Further, some bias may be introduced since the diagnostic sample will increase the person's level of awareness of vehicle problems beyond his usual level. What ever the technique used, it will be advisable to employ an experienced survey research organization to conduct this aspect of the data collection. Such an organization will have sufficient knowledge and experienced personnel to perform correctly this critical task.

An indirect measurement of the effect of the program on both vehicle condition and on owner maintenance practice can be obtained from reports of replacement part sales. This data will have to be collected on an ongoing basis from business firms in the test areas.

It cannot be captured from inventory records which either may be non-existent or may be structured in a way which makes data retrieval difficult. What must be done is to contract with firms to complete regular reports of their sales of parts which are frequently found defective at inspection. Ten outlets in each area should be selected from the following categories: automobile dealerships, gasoline service stations, automobile parts jobbers, and specialty repair facilities. Such data potentially can illustrate changes in repair activity following institution of an inspection system. It can be quite useful in the self-inspection and check lane system where little direct information on repairs will be available, but this data will be costly to obtain.

Furthermore, the samples will be strongly influenced by such external factors as changes in competitive position among different outlets, the extent of transient traffic through the area, seasonal variations, and general business conditions. Consequently, the priority assigned to this program will be quite low, and it can be dropped with little damage to the overall evaluation effort.

So far all the evaluation measures have dealt with intermediate criteria such as changes in mechanical condition and in owner behavior. An attempt can be made to measure the effects of the inspection activity on accident involvement. Police in the study counties will record the inspection certificate number and mileage of every vehicle involved in an accident on their standard accident report form. If these report forms are presently being coded for computer processing, it will entail little additional expense to add these items. From the inspection certificate number, the date and mileage at inspection can be ascertained. At the end of the inspection cycle, mileage information can be obtained on vehicles inspected at the same time as the accident-involved vehicles. The frequency of vehicles leaving the population can be also estimated in this way. Using the technique of cohort analysis, the probability of accident involvement as a function of time since inspection can be computed. The data can be further cross-classified on the basis of mileage, vehicle age, and vehicle type. Even without complicating factors, the exercise is not likely to produce significant results since the probability of a vehicle being involved in an accident in any short period of time is quite low. Unless the effect of the inspection effort is large, any change will be lost in random statistical fluctuations. The complicating factors, which reduce even further the likelihood of success being measured, include the multiplicity of accident producing elements, variations in mileage under different conditions, errors in accident reporting, out-of-area accidents, and seasonal variations. Nevertheless, the effort should be made since valuable information has occasionally been obtained from the most unpromising experiments (Reference 5-15).

A final note on evaluation deals with the composition of the evaluation team and the strategy that it should use in conducting the evaluation. At the head of the team should be a senior scientist with a large amount of experience in directing research programs involving the coordination of several efforts. His professional background can be in the social or physical sciences. At the second level, there will be one or two persons who together have a solid background in statistics and in assembling large data sets. At the third level will be two junior members with backgrounds in social science and engineering; they will have responsibility for executing much of the detail of the analysis. Much of the data gathering will be handled by persons not regularly associated with the evaluation team. These can be either part-time personnel, such as students, who work under the direct supervision of the evaluation team or personnel who have other regular duties, e.g., the check lane personnel when operating in the sampling mode. For the diagnostic sampling of vehicles and for the household surveys, specialized subcontractors will be retained. These data gathering efforts require the use of a large number of trained personnel for a relatively short period of time. Only organizations with on-going activities in this type of research can effectively handle the problem.

A good portion of the effort of the evaluation staff will be devoted to coordination of the various groups with which it will be working. They must have the full support of the sponsoring agency and of each operating agency in the program. Furthermore, operating personnel, with whom the evaluation group will work, must be thoroughly briefed and instructed on the role that they are to play in the program. Good cooperation at this level will greatly enhance the likelihood of the program's success. Conversely, if the operating personnel see the program as simply an imposition of extra duties upon them, they will not give it the complete cooperation needed for success.

5.7 Summary and Overview

A complete prepackaged program has not been presented. A number of elements, particularly the mechanics of specific inspection techniques, have been omitted. Some of the problems of experimental design have been handled by judgment, experience, and the use of standard statistical tables, rather than by rigorous analysis. To achieve the final program plan will require several man-years of effort drawing on a diverse set of skills. However, before the final plan can be produced, several key decisions must be made. The location and duration of the program will have to be settled, and the exact evaluation tools will have to be selected before a more detailed exercise will be useful.

What has been attempted is to outline in some detail a feasible experiment to test the relative effectiveness of three types of inspection programs. In this outline, attention has been given primarily to discussing the likely problems that will be encountered in establishing such an experiment and to recommending solutions to those problems. A set of goals has been established and a number of options to reach the goal have been described. By having these options so specified, the relevant decision makers can make an informed choice of the program which most fits their needs rather than being forced to accept one plan in toto.

Of the options, the recommended course is to use the three-in-one approach in a single state for five years. Evaluation will concentrate on measurement by the check lane sampling operation, street observations, interviews of persons in inspection, the owner survey, and accident data analysis.

REFERENCES

5-1 Michigan Department of State Police. Motor Vehicle Inspection Manual. East Lansing, Department of State Police, 1971.

5-2 Van Gorden, Dean. A Suggested Trial Substitute Periodic Motor Vehicle Inspection Program. Madison, Wisconsin, Office of Highway Safety Coordination, 1969.

5-3 American Standard Inspection Requirements for Motor Vehicles, Trailers, and Semitrailers Operated on Public Highways. (New York, New York): American Standards Association, 1963.

5-4 Report on the Motor Vehicle Inspection Program and Its Relationship to Highway Safety. (Trenton, N.J.): State of New Jersey Department of Law and Public Safety, 1963.

5-5 "Instructions for Official Inspection Stations." Boston, Massachusetts: Commonwealth of Massachusetts Department of Motor Vehicles, N.d.

5-6 O'Sullivan, Donal. Periodic Motor Vehicle Inspection in Ireland. An Appraisal and Work Program Outline. Dubline, Eire: Eire Department of Local Government, 1971. (Distributed in U.S. by Automobile Manufacturers Association.)

5-7 Boyd, A.S. Safety for Motor Vehicles in Use: A Report from the Secretary of Transportation. Washington, D.C.: Department of Transportation, June, 1968.

5-8 Used Car Safety Standards (Vehicle Safety Status). Final Report. St. Louis, Missouri: Automobile Club of Missouri, July, 1969.

5-9 Used Car Safety Standards. Final Report. Redondo Beach, California: TRW Systems Group, October, 1969.

5-10 Eisner, H., et al. An Investigation of Used Car Safety. Final Report. Silver Spring, Maryland: Operations Research, Incorporated, June, 1968.

5-11 Little, J.W. Motor Vehicle Inspection Administration: A Critical Review. Ann Arbor, Michigan: The University of Michigan, Highway Safety Research Institute, June, 1967.

5-12 Ibid. Motor Vehicle Inspection Legislation: A Survey of Patterns of Opposition. Ann Arbor, Michigan: The University of Michigan, Highway Safety Research Institute, January, 1967.

5-13 U.S. Senate, 91st Congress. Automobile Repair Industry: Hearings before the Subcommittee on Antitrust and Monopoly. Washington, D.C.: U.S. Government Printing Office, 1971.

5-14 Goodwin, W.I. Development of a Periodic Motor Vehicle Inspection System. Interim Report. Knoxville, Tennessee: University of Tennessee, Department of Civil Engineering, October, 1970.

5-15 Reinfert, Donald W. The University of North Carolina, Highway Safety Research Center. (Personal communication, July 19, 1971).

6.0 SYSTEM RESTORATION

There is little question that some form of emergency service is desirable for any community. In some highway safety program areas there is discussion as to whether the required capability need exist at all: for example, such a discussion continues in the field of periodic motor vehicle inspection. But the problem in emergency services is more one of determining the best level of operation than of deciding whether or not to have a capability at all.

From an experimental designer's point of view this complicates the problem; it would be much easier to measure the difference, say, between no ambulance service and a good ambulance service than to try to measure the (somewhat smaller) difference between two different training courses for ambulance attendants.

Nevertheless, that is the problem. Each community must make a conscious decision to implement the emergency service program it feels is best for itself, and the purpose of the present experiment is seen to be that of developing information which will be useful to the people who must make that choice.

Within the spirit of the Highway Safety Program Standards, a wide range of emergency service capability could be achieved. The minimum would be the formation of some planning committee, the enactment of legislation to require, say, the advance Red Cross course for all attendants, the purchase of a two-way radio system for dispatch, and use of modern--though not overly sophisticated--equipment. The maximum might employ fully trained paramedics, develop procedural manuals for debris handling, or aim toward some optimal allocation of resources to minimize response time, etc. We will describe these possible levels of service in more detail.

Three groups or kinds of activities which might be changes for the better have been identified in the area of restoration services. These stem mainly from existing requirements in the standards, and in fact constitute the elements for which some evidence of value is desired. They are: (1) training of personnel, (2) the acquisition of new (more capable) equipment, and (3) upgrading of the communications and management (control) aspects of the program. We have considered these as potentially independent changes, as paired changes (e.g., training plus new equipment), and as an aggregate change (all three together), and have concluded that for experimental (as well as practical) purposes, it would be best to improve all three simultaneously. The experimental cell structure shown in Figure 6-1 indicates the possible combinations. We recommend that only the lower level (or three-fold interaction cell) be implemented at three levels of intensity or complexity--a sort of good, better, best arrangement.

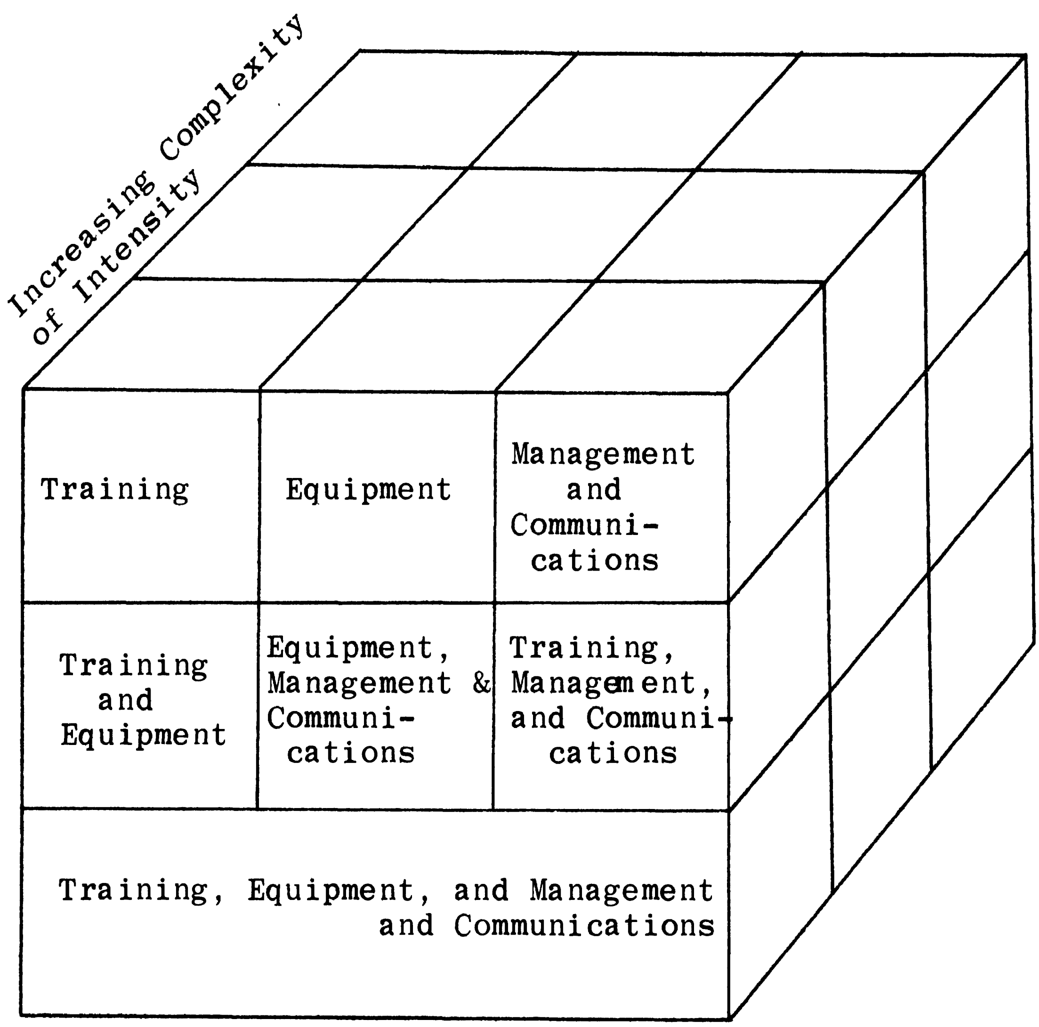


Figure 6-1

Experimental Program Matrix for the System Restoration Category

6.1 Evaluation of Recovery Systems

Recovery Systems have been defined as those elements in the highway traffic system which restore order after a crash--essentially the emergency medical and the debris removal activities. We have recommended that, for experimental purposes, both of these areas be changed (improved) in concert and that the effectiveness of the improvement be measured in each. In discussing evaluation procedures, however, we will concentrate first on the medical aspects, then on the problems of debris removal.

The evaluation of an emergency medical care system requires (as do other highway safety countermeasures) a careful selection and measurement of meaningful intermediate variables which are related to the ultimate objective--reduction of death and severe trauma by means of a correct emergency response. The evaluation design presented here provides a method for measuring both the quality (proper emergency diagnosis and treatment) and speed of response (time to scene of incident, time at scene, and time from scene to hospital). In addition, this design overcomes some of the difficult problems of combining the many complex and interrelated variables which effect the ultimate objective. This is done by using a qualified medical person--possibly an intern, a resident, or a skilled para-medical person--to gather information in a suitable report format on individual cases that resulted in death or slow recovery. These reports are presented to a jury of medical professionals who evaluate the action taken and make a judgement concerning its appropriateness. This jury would also be asked to determine if improved training, shorter response time, or better equipment would have resulted in a more favorable resolution of the emergency (e.g., prevention of death or reduction in the amount or length of subsequent treatment) and to order the importance of these factors.

The evaluation will be conducted by the comparison of three demonstration projects conducted at different levels of emergency medical resource allocation. The original concept had included a control area where no changes would be made, but where comparable measurements would be taken for comparison purposes. The control area requirement was rejected for two reasons: 1) the questionability of its value; and 2), the difficulty of obtaining a control area. Since before and after measurements are being made, the value of the control is to provide a baseline measure of change that occurred without experimental intervention (other than measurement). However, it is questionable whether the changes occurring in the control area really represent the changes in the project areas without the system improvements. The reason for skepticism is related to the anticipated problem of obtaining a good control area. It is doubtful that a local area would be willing to cooperate in measurement to the extent required, without some benefit accruing to the community.

In each of the three areas, before and after measurements of system performance would be made. The basic measure of project effectiveness would be a comparison of incremental changes in system performance over the three areas. Thus, for example, if the project which has the maximum allocation of resources truly results in an improvement, there should be a measurable change--which is of course related to system improvement--after the increase in medical resource allocation. It is further argued that the magnitude of this change should be greater than a similar measured change in the area where a smaller increase in emergency medical resources has occurred.

6.2 Time Schedule for Projects

A time schedule for the three demonstration programs is summarized in Figure 6-2. It is necessary to follow this rather long schedule in order to obtain a number of cases sufficient for adequate comparison. In addition, ample time must be provided to make changes to the system (e.g., conduct training and major equipment purchase) and to allow the changed system to stabilize prior to actually measuring its effectiveness. Immediately upon release of the contracts,

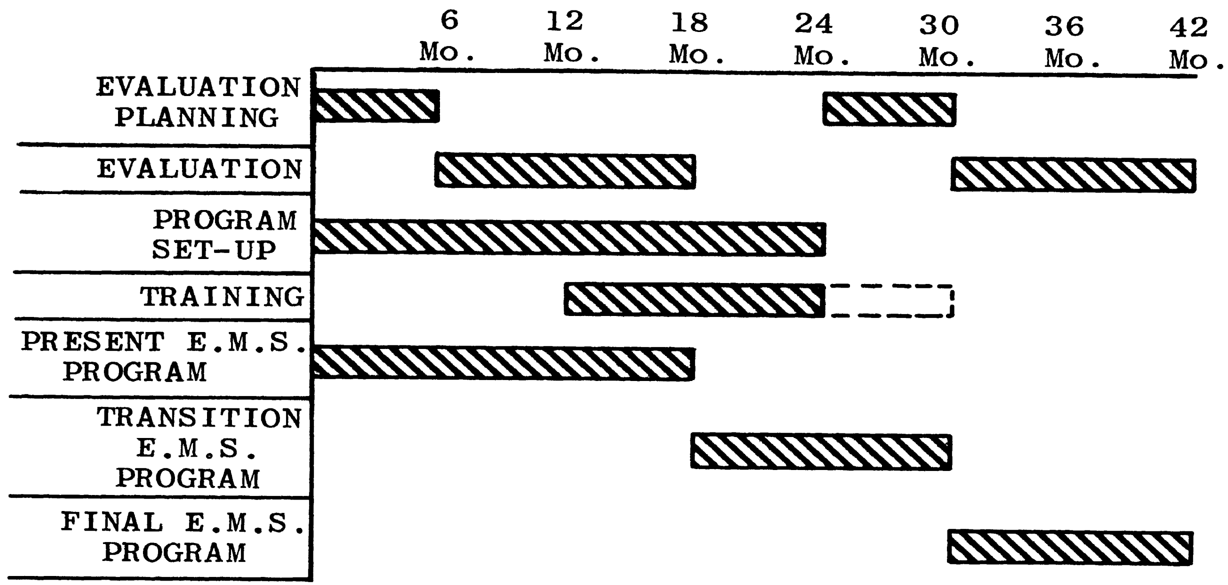


FIGURE 6-2. SCHEDULE FOR EMERGENCY MEDICAL ROAD SERVICE DEMONSTRATION PROJECTS

the local areas will have six months to set up the required data collection system and establish the medical jury. Data collection systems for the three projects will have to be coordinated to provide for a uniform level of data from each. It is also strongly recommended that the three medical juries meet jointly to establish procedures and criteria for judging the individual cases. By doing this it will be possible to compare the results under the best possible conditions.

During the first 18 months of the project all of the necessary arrangements will be made for the designated changes to the system. Thus at the beginning of the 19th month the proposed improvements should be implemented. It is planned that these changes can be implemented within a six month period. The system will then be allowed to operate for a period of six months in order to stabilize prior to actually beginning the formal evaluation. By conducting the project in this manner, the evaluation will compare the present system with an improved system that is operating under stable conditions. Thus, any of the problems and biases normally associated with start-up will not contaminate the results.

6.3 Description of Programs

As indicated previously, three separate demonstration projects will be conducted as the evaluation experiment for emergency medical care service. Each of the three areas chosen for these projects should be comparable in terms of population, geographic distribution of population, need for emergency medical service, and availability of regular medical care facilities. In addition, the pre-experiment level of emergency medical service should be approximately comparable. Another important criterion is of course a strong local interest in the improvement of the system. In particular, since local medical professionals will be asked to cooperate, their cooperation and support is essential. In this section we present a detailed discussion of the three alternative levels of service that should be compared.

6.3.1 The Basic Level of Service

This community will be expected to achieve a level of restoration service defined as minimal in the debris removal and emergency medical standards. Many states have not enacted specific training and licensing requirements, and it would be expected that the local community would provide such legislation if it is not covered by higher authority. A law similar to that in Michigan requiring each attendant or driver to take the advanced Red Cross Course, the Bureau of Mines Course, or an equivalent would be appropriate. It is expected that the community (county) would have an operating emergency medical committee whose function is to guide the improvement of the system from its original condition to the basic level.

Two way radio equipment should be included in the improved system, at least for the dispatch function, both for ambulance and for wreckers. But other than that, there should be no pressure for improvement. This community will serve as the base level for comparison with those using more complete training and equipment.

6.3.2 The Intermediate Level of Service

In this community the use of the NHTSA training course outlined in the Dunlap Associates reports (References 6-1 and 6-2) will be expected. This curriculum is general, but could be taught with particular emphasis on the highway emergency. Although the course need not be taught by physicians, incorporating it within the programs of the local medical society would provide several advantages. It would be much easier to include in-hospital training at a later date if this were done, and it would help establish rapport between the rescue workers and the hospital personnel. All of this would help to integrate the total emergency medical services within the community, and we believe that this would promote success within the program.

Two-way radio with central coordination and dispatch for the entire area is desired for this program, in order to reduce time delays. Dispatch could be handled by some governmental organization or by a private purveyor, but at least the medical aspects should be done by a single organization.

Ambulances should conform, as much as possible, to the NHTSA National Academy of Sciences recommendations reported in Reference 6-3. This document contains the considered opinion of a large number of people representing a variety of fields; the efficacy of the conclusions deserves to be fully tested. This program, with well trained attendents, should provide such an opportunity.

Systems Development Corporation, in their recent study of the highway cleanup problem conducted for NHTSA (Reference 6-4), recommended four training packages: (1) a brief orientation seminar for officials, (2) individual training packages for operational crews concerning hazardous materials identification, on-site safety procedures, extrication, etc., (3) a special course for control center personnel with emphasis on situation analysis, incident interrogation techniques, and (4) a concerted training program involving all, or some combination of agency personnel in simulated highway emergency situations to develop intra- and inter-agency coordination procedures. A program to develop a manual of local resources for debris removal is desired; this should be done during the first year. It will be expected that police and fire dispatchers throughout the jurisdiction will receive some periodic training, and the use of this manual will be incorporated in that training. It is suggested that the task of developing this manual be undertaken by one of the local law enforcement agencies, and that it be distributed throughout the region.

Finally a strong public information program is very desirable. This program should include publicity concerning the new medical service as well as phone numbers for emergency service. First aid training for the public might be included in this effort.

6.3.3 The Top Level of Service

The third community should provide a level of emergency service second to none. It will be used experimentally to show the value of a relatively expensive capability adequate to handle most expected emergencies.

The ambulance attendant training should be equivalent to that received by a military medical corpsman, preferably with additional civilian training in connection with the local hospital facilities. Although current statutes generally prohibit these attendants from performing certain duties (such as endotracheal intubation or starting intravenous fluids), it is proposed that they be trained in these tasks, and that the community be encouraged to pass enabling legislation. In any case, ambulances should be equipped with equipment necessary to perform tasks of this nature on the chance that there may often be a physician present at the scene of an emergency. The basic ambulance should be the same as that discussed earlier, but the use of additional equipment such as telemetry for vital sign data should be encouraged.

Communication should include two way radio for dispatch, and should also include the equivalent of the HEAR (Hospital Emergency Administrative Radio) system permitting direct conversation between the ambulance personnel and the hospital emergency rooms. The professional status of the attendants, and the use of this radio is expected to increase the rapport between ambulance crews and the medical community. Cooperation of the local physicians in connection with the hospital-to-ambulance radio system should be aggressively sought.

Debris removal activities should be about the same as in the intermediate level of service, except that the introduction of special equipment into this area should be encouraged. Systems Development Corporation (Reference 6-4) has recommended the introduction of several new kinds of equipment which may be developed over the next several years. If they become available in time, this program would provide a useful test bed. This equipment includes push bumpers on all emergency vehicles (which could be implemented at any time), sweepers, cranes, vacuum pickups, and multipurpose vehicles possessing the capabilities of hoists and fork-lifts at the same time. This equipment, in concert with trained personnel and adequate communications capability, should get a good test in this community.

6.4 Measures of Effectiveness

The ultimate success of an emergency medical care system is of course defined by a reduction in the occurrence of death, permanent disability and long term treatment among the patients served. However, these are related to a number of factors that are independent of the emergency medical care system: for example, the severity of the initial incident, the age and health of the patient prior to the incident, and subsequent medical care after the emergency. Thus a mere tabulation of "successes" and "failures" is not sufficient to measure program effectiveness.

An analysis of emergency medical care reveals that improvements should be reflected in terms of the following factors:

1. Reduction in service time defined as
    a. Time from the occurrence of the incident (e.g., a highway crash)
    b. Time at scene of incident
    c. Time from scene to hospital
2. Correct immediate diagnosis of the problem
3. Correct emergency treatment of the problem

For example, more ambulances and better communication should result in reduced service times. Better training should result in improved diagnosis and correct treatment at the scene in a higher percentage of cases.

These factors are related by a complex interactive process to the ultimate criteria listed above. For example, the relationship between the value of reduced service time and correct treatment depends on factors unique to a particular case.

The development of a predictive evaluation model that properly relates the contribution of the above factors to the ultimate objective is a difficult if not impossible task. We are proposing instead the use of available medical expertise to make post facto classification judgements, employing all information concerning the incident. Since precise data is a pre-requisite to this evaluation process, we propose that a single qualified medical person be in charge of data collection. This person will collect all pertinent data on a case-by-case basis. He will then prepare a report for the emergency medical service jury. This jury (composed of doctors and possibly other knowledgeable health professionals) will evaluate each case to determine at least the following:

1. Was the attendant's diagnosis correct? If it was not:
   a. Is it possible that a reasonable improvement in the training of ambulance attendants could have resulted in a correct diagnosis?
   b. Did the incorrect diagnosis have a detrimental effect on the patient?
2. Was the attendant's treatment correct? If it was not:
   a. Is it possible that a reasonable improvement in the training of ambulance attendants could have resulted in a correct treatment?
   b. Did the incorrect treatment have a detrimental effect on the patient?
3. a. Was the total service time within the range normally experienced for this ambulance system?
   b. Would a reduction in total service time have helped the patient? How large a reduction in time would have been required to accomplish this?
4. a. Was the elapsed time from the occurrence of the incident to the arrival of the ambulance within the range normally experienced for this ambulance system?
   b. Would a reduction in this time have helped the patient? How large a reduction in time would have been required?

5. a. Did the ambulance attendants make a decision either directly or implicitly through their diagnosis and subsequent treatment--to sacrifice either treatment or total service time in favor of the other?

   b. Which of the two actions did they favor?

   c. Was this judgement made correctly?

Questions 1 and 2 above deal with the training and equipment provided in the system. If better training and equipment do not result in an increase in the number of correct diagnoses and treatments then it is questionable whether or not these improvements are worth their added cost. Similarly, improved location and dispatching of ambulances should result in a reduction in the various service times. Even if these times are reduced, it is still necessary to determine the effect of this reduction on the patient's health.

Approximately 300 cases per site per year will be utilized; for each a detailed case history will be prepared. These cases will be selected by severity, and at this writing the suggested rule for selection is that all cases involving broken bones and/or requiring admission to the hospital be used. The case history will contain a description of the conditions of the accident and injury, a tabulation of the treatment attempted or given by the ambulance personnel, and an assessment of the case by the participating physician detailing the cause of death or long term disability, and emphasizing recovery time. For each case history a team of physicians (the medical jury) will be asked to complete the forms shown in Figure 6-3. This data will be used subsequently to determine the relative value of enhanced treatment capability and reduced service times associated with the different sites.

The data from these forms will be entered into a file for analysis using one entry for each case. The data will be of the following form:

Variable #1	Name of site
Variable #2	Before or after period

Figure 6-3

Medical Jury Case History Questionnaire

Was the diagnosis given by the ambulance attendent:
(circle one) Correct Incorrect

If incorrect, which level of training would most likely have been necessary to insure correct diagnosis? (circle one)

Basic Red Cross	Advanced Red Cross
NHTSA	Military Medical Corpsman
General Practitioner	Surgeon

In this case the times associated with the rescue process were as follows: Circle the appropriate column if you believe that a shorter time would have contributed materially to the recovery process.

Time from injury occurrence to arrival of ambulance scene________.	YES	NO
Time at scene ________.	YES	NO
Time from scene to hospital ________.	YES	NO

If the answer to any of the above is yes, please explain what amount of time saved would have been helpful and why.

____

____

____

	Control Light Bleeding	Control Severe Bleeding	Guarantee Adequate Respir.	Splint Long Bones	Use of Back Board	Cardio-Pulmonary Resus.
Did this case require treatment at the scene?						
Was the treatment provided by the ambulance attendant:						
Correct and Successful?						
Correct and Unsuccessful?						
Indicated but not attempted?						
Incorrect?						
Training level * indicated						

*If the treatment attempted was incorrect or unsuccessful, please indicate in the appropriate cell above which level of training (for the attendant) would most likely have been the minimum necessary to insure correct treatment.

1 - Basic Red Cross 2 - Advanced Red Cross 3 - NHTSA

4 - Military Medical Corpsman 4 - General Practitioner 5 - Surgeon

6 - None of the above

FIGURE 6-3. (CONTINUED)

Variable #3	Date (year, month, day)
Variable #4	Patient sex (male, female)
Variable #5	Patient age (in years at time of accident)
Variable #6	Patient's general health at time of accident (poor, fair, good, excellent)
Variable #7-48	Matrix of Figure 6-3
Variable #49-55	Responses to diagnosis question
Variable #56-61	Responses to time question, including required times in minutes.

Approximately 1800 cases will exist in this form at the end of the experimental program, and they may be analyzed in a number of ways. The particular analyses visualized at this time include a before and after measurement of the percentage of correct diagnoses, the percentage of correct treatments, and a comparison of the percentage of cases in which time was felt to be of importance.

The continuous scale for time, and the categorical scale for training level are intended to permit a comparison of the relative value of these two system characteristics.

The size of the experimental program, as measured by the number of cases for which data will be collected, depends mainly on two factors; how large a difference we hope to achieve (e.g., the percentage of correct diagnoses by ambulance attendants has increased from 70% to 82%); and what assurance we require that the above kind of statement is correct. Usually the latter problem is solved by choosing some acceptable level of statistical significance. A common level of assurance (among statisticians at least) is 95% meaning that if we decide from the experimental evidence that things are truly better we will be correct in this interpretation all but one out of 20 times. It is unfortunate that such complex conditional statements must be made with respect to what must seem like straightforward evaluation problems, but this is the language of statistics. At any rate, if we agree that we will accept 95% assurance that our results were not obtained by chance, then we are in a position to consider the trade-off between the size of the experiment (number of cases examined) and the minimum differences we can detect.

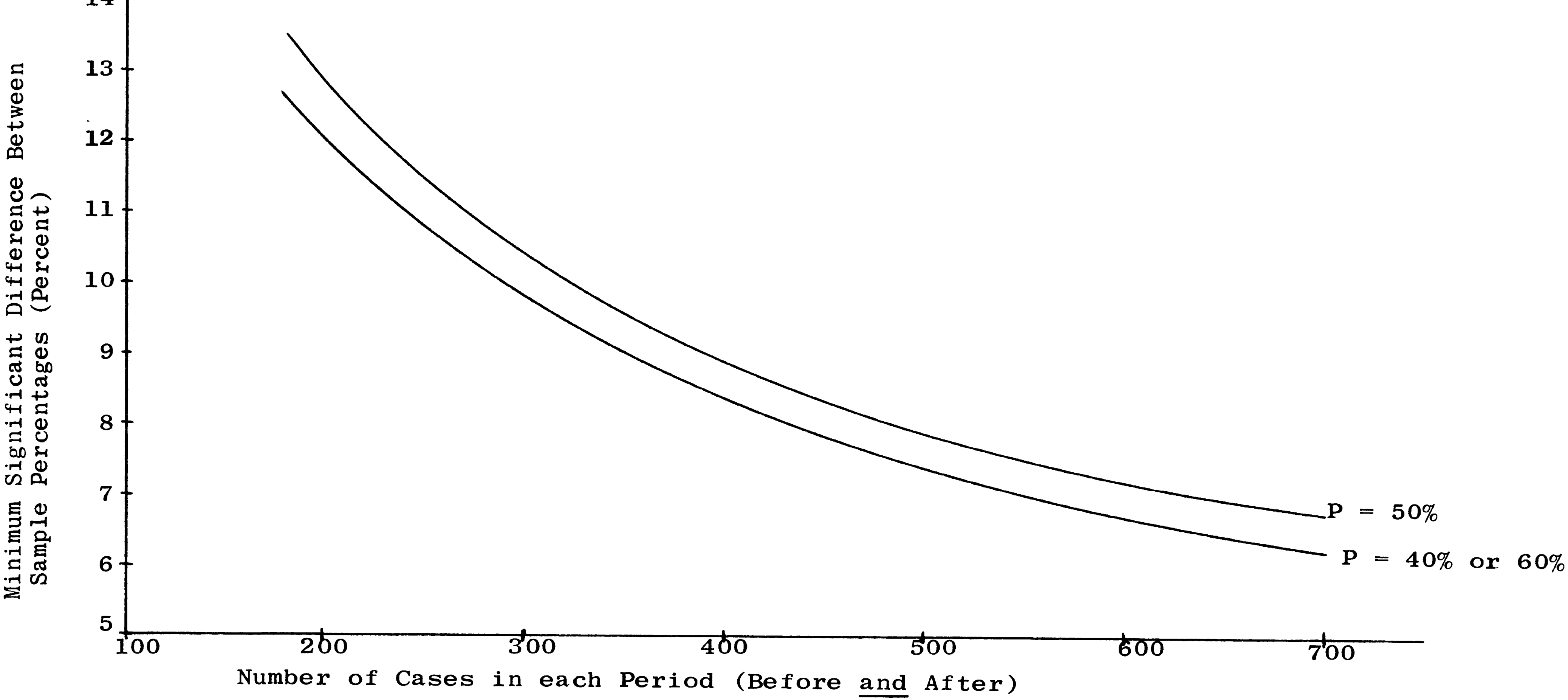
Figure 6-4 shows this tradeoff applied to the present problem. (See Appendix C for derivation assumptions and procedures). Suppose that 50% of the cases observed in the "before" period had an error in diagnosis, and that 40% were in error in the "after" period. Enter the graph with the number of observations (suppose there are 200 cases), proceed up to the intersection with the 50% curve, and then read the percentage--in this case 12.7%. We conclude that the observed reduction was not enough to assure us that the "after" system was really better. Alternatively, we may wish to know that if we observe a 10% difference we must have a certain number of cases to look at. This indicates that 320 cases are required.

It should be obvious that all of these numbers are not known in advance of the program. Nevertheless, in a previous study of emergency medical care activities in the city of Detroit we found that between 40 and 60 percent of the serious cases handled contained an error that could have been avoided by better training and/or equipment. Thus we anticipate that the pre-program evaluation of serious cases by the medical jury will result in a percentage of errors somewhere near that range. However, even if the percentage is not between 40 and 60, the relative tradeoff between number of cases and minimum significant difference will be the same. Thus the use of Figure 6-4 will result in the same number of cases being judged regardless of the base percentage of errors. A difference in the base percentage of cases with errors will only effect the minimum significant difference.

We have somewhat arbitrarily selected a goal of 300 to 400 cases before and 300 to 400 cases after the change to provide an adequate trade-off between program size and sensitivity to change. Thus, criteria for choosing program sites should include the requirement that a sufficient number of possible cases occur each year.

Figure 6-4

Comparison of the Number of Cases Submitted to the Medical Jury, Before and After Program Changes - vs. the Minimum Significant Difference



6.4.1 The Medical Jury and Its Information Gathering Function

The success or failure of an emergency medical system is measured by a number of intangible factors that are difficult if not impossible to quantify in a systematic manner. Death of a particular patient may result from improper diagnosis, improper treatment, or delays in emergency treatment. However, it may merely be the result of very severe injuries which any reasonable system could not be expected to deal with successfully. In an attempt to provide some objectivity in the evaluation, we propose that competent judges of medical care be made an integral part of the measurement system, to consider each case subjectively. In addition, we recommend that a competent medical person be employed to gather information that the medical judges require and to document the reasons for making decisions as to the correct handling of particular cases.

The use of a local team of medical professionals could lead to occasional biased conclusions because of preconceived opinions about the particular system and the people within it. To prevent this, the team must understand the importance of objective evaluation in the proper choice of medical treatment systems. In addition, randomly selected cases from each program should be reviewed by the program monitor to compare the quality of the judgements between the various projects.

We believe that it is very important that one person be assigned the task of collecting the data for all cases to be evaluated by the medical jury. This person will have to do a great deal of investigatory work such as interviewing patients, ambulance attendants, police officers and others. In addition he will have to review existing medical diagnosis and treatment reports and interview doctors to obtain additional information concerning specific cases. Because of the nature of the job, the person chosen should have medical training; an intern, medical resident or senior para-medical specialist would be a sound choice here. This person will have to understand medical terminology and treatment rationale in order to provide a complete description of the treatment cycle and potential shortcomings. In addition, he will have to establish a number of informal communication links with various doctors, nurses, and others associated with the emergency and normal treatment of patients.

The medical jury should not rely on routine reports, either those presently available or new ones superimposed on the system. These reports may not be complete enough to permit comparison between cases. Examination of the records after the fact leaves unanswered such questions as: Is this an indication that a particular condition did not exist? or is it merely a result of the information not being ascertained? Situations like these render reports almost useless for purposes of rigid comparisons between the treatment received and a potential alternative. This problem can only be overcome by a competent medical person contributing the detail to each report in a consistent manner.

Individual projects may want to develop procedures and forms which will help them to obtain the data discussed above in a complete and straightforward manner. However, we have included sample copies of data collection forms that have been field tested and modified in other emergency medical care projects. It may also be possible to develop additional routine methods for gathering some items of information. For example, in Washtenaw County the emergency system ambulance attendants are required to report by radio the time that they leave, and the time that they arrive at the hospital. This information is recorded by means of a time stamp clock on a card describing the particular run. This system has proven to be more accurate than the use of police or ambulance attendant's hand-recorded reports which tend to be rounded to the nearest five minutes. Since one or two minutes of time saved might be critical, it is important to have a great deal of accuracy in this phase of the data collection.

We have made a number of specific recommendations concerning data collection procedures and formats. There are undoubtedly others that will result from a systematic analysis of the particular areas in which the demonstration projects are to be conducted. During the six month planning period, detailed procedures should be established. A high degree of commonality among the project areas should be sought in the planning process.

The creation of a good data collection system will not in any way replace the case investigator. There are a number of subjective items critical for evaluation that can only be obtained by direct contact with the medical professionals involved in the treatment. This subjective information must be structured in a consistent manner in order to provide for comparison between cases.

6.4.2 Measurements Before the Application of System Improvements

Since the problems faced by emergency medical care facilities vary from one area to another, it is not reasonable to merely compare the absolute level of program effectiveness at different sites. For example, the fact that one area has a smaller percentage of incorrect diagnoses and/or treatments may result from less severe incidents and/or a better level of hospital care. The important question when resources are being allocated to a particular program is: How have these additional resources affected that program? It is certainly important that each experimental area have adequate hospital facilities. However, when evaluating the emergency medical care system, it is desirable and perhaps necessary to separate out the contribution of the hospital to overall patient recovery. Therefore it is imporant that all measurements be made on a before and after basis and that comparisons between programs be made in terms of the magnitude of the change.

Since the evaluation requires measurements before the system is improved it will be necessary to have the medical jury system operational for a period of time prior to the implementation of the change. It is recommended that complete data be collected for a period of one year before the system is changed. By making the same measurements after this baseline period there will be a rational method for determining if changes have actually occurred.

Collecting data prior to the changes may cause some difficulties. This is especially true since the measurements depend on the use of a local team of medical experts. These people may not believe that this is a fruitful use of their time. Therefore it is important that they be convinced of its importance. As an added incentive, the program should be designed in such a manner that recommendations from the medical jury can be fed into the design of the system that is implemented in the particular area. Of course their recommendations will have to be somewhat limited, since it may be difficult to obtain an unbiased measurement of system components that are the "pet" projects of those persons doing the measurement. These are important considerations and are deficiencies in the use of a medical jury for evaluation. However, the alternative of not using some human expertise to integrate the raw measurements has even more disadvantages. We believe that this design is a suitable compromise. The problems relate to human personalities and their interaction; thus it is difficult to present general solutions and we will not attempt to do so. These problems are not insurmountable; however, they will take some time to deal with. This is one of the reasons that we originally set community acceptance as an important criteria for the choosing of areas in which to establish programs.

6.4.3 Evaluation of the Debris Removal Activities

The effect of inadequate debris removal is evidenced in several ways. In Oakland County, Michigan approximately 0.3% of some 30,000 accidents had "a previous accident" listed as a contributing cause. Our experience has shown that often the second accident occurs immediately (say a few seconds) after the first, and it is not clear that more rapid removal of the first accident would have made the situation much better. Nevertheless, the frequency of second accidents at a later time is not infrequent. State, city and county police alike have complained about the frequency that their accident investigating cars are struck, even when their emergency flasher lights are operating. So it is possible that a change in this subset of the accident data may appear if a real effort is made to minimize the time that debris of wrecked vehicles remains on the road. Accident data for the communities involved will be searched for this factor, and a bi-level form for reporting such accidents is suggested for all communities participating. The raw frequencies of such incidents will be compared in the before and after periods.

For the really obstructive debris, however, it is not possible to do much statistical analysis. Rather a sort of case study approach is recommended. For each community participating, any instances in which there is spillage of a large load (a steel truck, a brick truck), leakage of some contaminating fluid (gasoline, etc.), any radioactive accidents or explosions, and any cases of extrication when special equipment must be used, will be reported in detail, and a subjective judgment made as to the adequacy of the methods and equipment in use. Desirable as a statistically study might be, there will not be enough cases in all of the programs discussed here to perform one; and we believe that the anecdotal method will provide cogent arguments as to the value of specialized equipment and training if they are indeed useful.

6.5 Cost of the Restoration Programs

Costs will depend in part on the particular wage situation in the region selected for demonstration programs, and in part on the training and equipment level achieved. Ambulance system operations have been reported which average from a few dollars per run to well over 100 dollars per run--this being more a function of the load per unit than anything else.

In each area chosen for this effort, it is expected that somewhere between 4 and 8 ambulances will provide adequate coverage. Volz (Reference 6-5), in a simulation of the ambulance allocation problem found five or six ambulances as a likely optimal solution for a county of 750 square miles and 200,000 residents. Initial vehicle cost, including communications equipment, should range from about $7,000 to $16,000 each. Wages, on the other hand, may vary considerably from a basic minimum wage for the least service to perhaps $10,000 or more per year per attendant for the best service. At $2.00 per attendant hour, and two attendants per vehicle the total wages per ambulance would be about $40,000 per year. But a more professional attendant (at say $5.00 per hour) as is recommended in the higher level program would take the personnel costs (per ambulance) to about $100,000. If full subsidy of the three programs is required, the average operating costs for each site are estimated at $500,000 per year (for five ambulances), including necessary dispatching, etc. As has been discussed, the improved ambulance service may only involve some change in an existing service, and many of these costs be covered by existing budgets. It is conceivable that a demonstration program could be run using all volunteer agencies, or an acceptable program might involve part time use of fire department personnel. This would be most possible for the first or second levels of service, and costs then might be much less. The major additional effort in the debris removal field is the preparation of a handbook in the middle and highest level programs. This might best be accomplished by a junior employee of a traffic engineering department, and we have estimated six man months for that. Total estimated costs for each site are given in Table 6-1.

The problems of evaluation will require that some one person follow the several programs on a continuing basis--keeping data centrally and coordinating the activities in three sites. During certain periods he may require assistance in the preparation of data. At each site, however, there must also be a coordinating physician we estimate 25% of his time over the period--and at least one assistant to monitor the data collection and generally assist in the

Table 6-1

$3\frac{1}{2}$ year Costs, Restoration Demonstration Programs

	Site I	Site II	Site III	Program Management
Ambulance equipment (5 vehicles)	$35,000	$60,000	$80,000	
Personnel for 5 vehicles	@ $2.00/hr $700,000	@ $3.50/hr $1,325,000	@ $5.00/hr $1,750,000	
Planning, (committees, debris removal handbook, etc.)	$10,000	$20,000	$20,000	
Monitoring data taking, research activities	$75,000	$75,000	$75,000	$100,000
Training programs	----	$5,000	$5,000	
Total	$820,000	$1,485,000	$1,930,000	$100,000
GRAND TOTAL		$4,335,000		

local operations. In the sites with debris removal activity we suggest that some local government employee be retained to monitor and record particular debris removal operations.

Ambulance systems, while of great potential value to the highway injury problem, can hardly be used exclusively for that function. Experience in the county used as a base for this study indicated that there were about 1300 emergency runs from automobile crash scenes to a hospital each year. But the total number of ambulance runs (including many non-emergency carries, such as hospital transfers) was on the order of 6000. In the most elaborate program we have described, the average cost of an ambulance run, then, will be about $107. For the lowest level treatment this will be $45.00.

All of these costs may seem too high to a local governmental unit, and it may be difficult to convince that unit to continue service after some federal funding disappears. A currently acceptable life value, computed from future earnings, is $140,000; the range of costs above thus span from 2 to 5 lives per year for a cost benefit ratio of one.

It is nearly certain that the difference among these ambulance services will not yield statistically meaningful figures regarding the number of lives saved. There may, of course, be anecdotal evidence that several people were revived by one system who might otherwise have died; but it is almost as likely that this would be the poorer system--and this would be a disappointing result.

Nevertheless, we believe that a carefully controlled experiment, resulting in the called for measures of performance and their subsequent analysis by a medical team will provide information which a city or county government can use to decide which level of emergency service is most appropriate for its needs.

REFERENCES

6-1 Basic Training Program for Emergency Medical Technicians--Ambulance Concepts and Recommendations, October 1969, Dunlap and Associates, Inc.

6-2 Basic Training Program for Emergency Medical Technicians--Ambulance Course Guide and Course Coordinator Orientation Program, October 1969, Dunlap and Associates, Inc.

6-3 Medical Requirements for Ambulance Design and Equipment, September, 1968, Committee on Medical Services, Division of Medical Sciences, National Academy of Sciences, National Research Council.

6-4 Highway Debris Hazard Control and Cleanup Study, March, 1971, Systems Development Corp., DOT Contract FH-11-7274.

6-5 Optimum Ambulance Location in Semi-Rural Areas, R.A. Volz, Transportation Science, Vol. 5, No. 2, May 1971.

7.0 HIGHWAY REGULATION

Program efforts in the Highway Regulation Category are directed toward the evaluation of countermeasures that affect physical characteristics of the road and its environment. This includes changes related to highway geometrics, to traffic control devices, and to that portion of the pedestrian - highway interface that deals with equipment or pedestrian control devices (ie., cross walks, control lights, etc.).

Before detailing the program plans for this category, however, a brief review of the current state of development in Highway Regulation will be presented.

7.1 Review of the Highway Regulation Category

As indicated by the division of tasks used above, there are three Highway Safety Program Standards applicable to the Highway Regulation category:

1) Highway Design, Construction, and Maintenance
2) Traffic Control Devices
3) Pedestrian Safety

For the sake of completeness, therefore, we will review the requirements that these standards impose, and the activities that they have generated in the safety community.

7.1.1 Highway Safety Program Standard Requirements

Due to the nature of the tasks associated with the Highway Regulation Category, the Safety Standard requirements tend to be quite specific in their content when compared to one of the other five categories. For example, the Codes and Laws Standard calls for "unification of our system of traffic control regulations" while the Highway Design Standard requires for example "breakaway sign and lighting supports." The latter is obviously a much more straightforward task that is less likely to be interpreted in various ways.

The purpose of the Highway Design Standard is "To insure: (a) that existing streets and highways are maintained in a condition that promotes safety, (b) that capital improvements either to modernize existing roads or to provide new facilities meet approved safety standards, and (c) that appropriate precautions are taken to protect passing motorists as well as highway workers from accident involvement at highway construction sites." To accomplish this there are a number of requirements relating to specific design considerations (lighting, skid resistant pavement, etc.)

The aim of the Traffic Control Devices Standard is to promote the full use of existing knowledge in the field and the uniform application of existing control devices. The major requirements of the Standard consequently involve the upgrading of existing devices to conform with Federal Highway Standards.

The Pedestrian Safety Standard recognizes the pedestrian as an integral part of the highway traffic system and consequently requires the construction of facilities for the safe passage of pedestrians in the system (signs, land use planning, bridges, lights).

7.1.2 Review of Past Work in Each Area

Four principal activities in the Pedestrian Safety field have been funded under "402" sponsorship. They are:

1) Studies and/or implementations of new school crossings
2) Development of safety curriculums (mostly at the state level)
3) The construction of "safety towns" -- actually training aids in the form of model towns to train children as pedestrians
4) The establishment of school crossing guard facilities.

From the literature, a few more activities are evident:

1) Manuals for pedestrian crossing design
2) Discussions of pedestrian malls or pedestrian pathways in urban areas
3) Descriptions of special techniques such as zebra crossings, overheads, subways, and talking traffic lights
4) The organization and operation of school patrols

The status of knowledge in the pedestrian field is good and is generally accepted by the public agencies. Emphasis in the literature as well as in active programs has been on the issues of protection for the elderly and for school children, problem identification, and training.

In the Traffic Control Devices field, the majority of work has not been concerned with modification of the traffic system but has dealt almost entirely with surveys or inventories of existing sign and signal conditions in anticipation of future improvements. This condition has resulted from the current restrictions on hardware spending. The literature, on the other hand, describes a number of activities that may be condensed into the following groups:

1) Computer control of signals with optimum timing
2) Pedestrian crossings and signals
3) Rumble strips and stripes
4) Roadway delineation
5) Freeway merging control
6) Symbolic signing
7) Illumination of signs

Many of these activities have been small scale test programs and evaluations while others are simply discussions of applicable techniques.

In the Highway Design area the "402" programs are again primarily concerned with the study or inventory of existing conditions. The larger programs include a Nebraska study to correct the design of railroad crossings; another Nebraska program to inventory the skid resistance characteristics of several pavement types; a North Dakota inventory system for cataloging roadway and bridge hazards; and efforts in California to correct railroad crossing deficiencies.

Again, however, the literature abounds with ideas including many reports of innovative operational experience. In addition to some emphasis on safety in the general design literature itself, there are reports concerning the specifics of guard rails median barriers, breakaway signs, anti-glare fences, railroad grade crossings, bridge rails, and pavement characteristics.

There are many things that need to be done in the Highway Design area to achieve conformance with the standard. Indeed, some new roads are still being constructed with inherent major safety deficiencies such as rigid sign supports, improper bridge rails, and inadequate lighting. Yet the fault does not seem to result from the lack of modern safety techniques so much as in the failure of governmental agencies to revise their programming, funding, and contracting procedures. In short the situation seems to need an education program based on present knowledge of highway safety administration rather than a demonstration of feasible countermeasures.

7.2 Experiment Design Rationale

A number of factors affect the design of an evaluation experiment in the Highway Regulation Category; among the most important are:

1) The current state of knowledge in countermeasure development.
2) The current state of countermeasure deployment.
3) The degree to which safety practitioners accept the current state-of-the-art knowledge.
4) The quality of existing evaluation and information dissemination functions.
5) The diversity of specific countermeasures that are applicable.

In the review presented above, it was indicated that the current state of knowledge in countermeasure development is relatively good. The suggested techniques and equipment to be employed certainly come no where near full utilization of current technological capabilities, but there is active consideration of a number of techniques more advanced than any currently employed. Thus, while the status of countermeasures could be improved through application of technology, it seems adequate for the present state of deployment.

Because the Highway Design field is a well established engineering discipline, there are usually evaluation efforts carried out to determine the utility of new or existing countermeasures. While some of these efforts leave much to be desired in the way of accurate and definitive results, in general many of the evaluations are adequately performed and add to the existing body of knowledge. Evaluation efforts in the Highway Design category are probably at a higher level than any other category. The dissemination of information to the safety community at large is also good since reports and journals such as the Highway Research Record regularly report the results of investigations together with critical comments regarding their value. This exposure is extremely useful in weeding out unsubstantiated conclusions and in suggesting further areas for improvement.

The real difficulty in Highway Regulation appears to be the current state of countermeasures deployment. As mentioned earlier, there are many concepts whose value is undoubted, such as breakaway sign supports and light posts. Yet such countermeasures seem to take on inordinate amount of time to get into widespread use. This situation is particularly deplorable in highway safety since ones' mistakes are not shot into outer space but are cast in concrete and steel. This state of affairs puts a severe crimp in the utility of a countermeasure demonstration program whose objective is to convince people that certain selected safety countermeasures are (or are not) worthwhile and should be implemented. In effect people are convinced that the measures are worthwhile: the machinery has simply failed to act on the conviction, because of economic or political factors.

A more practical problem in the design of a countermeasure evaluation program is the diversity of specific techniques applicable to the highway category. For example, the use of roadway lighting, skid-resistant pavement, and traffic regulation at construction sites is called for by the Highway Design standard. In terms of the basic "building block" program plan presented in Volume 2 (see, for example, Figure 3, Volume 2), this means a large number of specific countermeasure "groups" with littler interaction among them. As a further detracting factor, it is highly possible that a single site location to carry out a number of these tasks simultaneously would be difficult to find.

From an evaluation of all the factors presented above it is our conclusion that a full countermeasure demonstration program in the Highway Regulation Program Category is not a feasible undertaking at the present time. This conclusion must not be misconstrued to imply that nothing need be done, however. Our concern is simply that what does need to be done cannot be accomplished by a single, large-scale, demonstration-type program. In light of the disparity between the state of knowledge in the field and the practical implementation of this knowledge, an administrative effort seems to be the most useful alternative at this time. In parallel with the treatment given to the other program categories it is suggested that the master building block program plan be used as a long term planning guide for countermeasure evaluation and implementation and that the existing efforts being carried out autonomously be encouraged to combine under this plan to provide a needed degree of cohesion for the overall activities.

A basic premise of this study is that the design and evaluation of experiments is as important at the current state of highway safety implementation as the techniques that are to be evaluated. Consequently, a program plan in the highway area will be discussed below to indicate the planning procedures recommended and some or the evaluation techniques that are applicable. This plan is put forth as a prototype effort; it is a suggested effort which could very well occupy one of the blocks in the master plan. From the review of past accomplishments, an experiment in freeway signing has been selected for discussion. There is considerable current interest in this field, and a significant amount of effort has been expended in developing measures of effectiveness for highway modifications. Consequently this is a timely subject to use as an illustration of our planning approach.

7.3 A Beltway Signing Experiment

According to the AASHO Interstate signing manual:

> "The design for signs of the Interstate system must be approached on the premise that the signing is primarily for the benefit of drivers who are not familiar with the route or the area. Signs must contain messages appropriate to the needs of these drivers. The sign legend must be carefully selected and designed for easy reading and the signs themselves must be priminently and effectively displayed at the proper location so that drivers will tend to react promptly, naturally, and safely to the traffic and design conditions encountered."

In spite of this basic directive, an NCHRP study published in 1969 found that (Reference 7-1):

> "Information needs for the motorist present a problem; it is found that the motorist considers official signing his chief annoyance when travelling. Also, signs for services, according to the motorists' indicated preferences, are not nearly sufficient."

While this conclusion is primarily concerned with the service aspects of signing, other studies have shown the need for adequate directional signing in highway and freeway environments (References 7-2 through 7-6). Beltways in particular pose a difficult orientation problem for the unfamiliar motorist (and consequently the sign designer) because the usual destination signs typically used on interstate freeways to indicate location or direction have limited significance on a beltway.

7.3.1 Program Goals

Ultimate goals for the applied countermeasures are the reduction of injuries and fatalities through the use of improved freeway signing. Typically these goals are much too broad and non-specific to serve as guidelines in program planning and intermediate goals must be defined.

In the past few years, the use of symbolic signing has been investigated as a technique to transmit needed information to the motorist more effectively than is possible through conventional legend signing (eg., Reference 7-4). Indeed, the 1970 version of the Uniform Traffic Control Devices manual shows an increased usage of symbology reflecting this trend. An implementable goal for a signing experiment is then to determine the relative utility of symbolic and legend signs for effective information transfer at troublesome freeway exits. This determination could include the types of information best passed by each sign and how this is related to the environmental surroundings.

A number of evaluation techniques have also been tested in recent years in an attempt to relate the countermeasure effect directly to measureable modifications of the traffic pattern in the vicinity of the countermeasure application (References 7-8 through 7-13). The "Traffic Conflict" technique is the best known of these attempts. Although these attempts have shown the capability to gauge accident potential in a relatively short period of time, they have not been successful in predicting ultimate effects (fatalities, etc.). As we have discussed throughout this report, this is due to the fact that ultimate criteria are affected by a great number of complex, little-understood factors in addition to the effect produced by the countermeasure. A second worthwhile program goal, therefore, is the use of traffic conflicts (or other such intermediate criteria) in a carefully instrumented environment to determine the factors that influence the relationship between conflicts, accident potential, and ultimate criteria. That is, what (generally unmeasured) environmental and social factors affect the relationship between conflict measurement and accidents, and how great is the effect.

7.3.2 The General Program Plan

The basic prototype program plan, in compliance with the goals defined above, employs a multi-exit beltway surrounding a large city to determine the relative efficacy of symbolic signing in informing drivers of critical information. Beltways around cities carry a significant amount of out-of-town traffic on a year around basis and thus provide a suitable experimental setting because:

1) Many users are unfamiliar with the system;
2) There are many closely spaced decision points;
3) There may be problems assocated with orientation;
4) There are frequent junctions with major routes;
5) Traffic volume is usually heavy.

On a given beltway, then, a number of troublesome exits can be chosen for treatment, while the remaining exits are left in their pre-experiment state. The nature of the signing used at each of the chosen exits should be in accordance with the latest AASHO standards and should be adaptively designed by a qualified engineering team to solve the suspected problems at each location.

The existence of many decision points on a beltway makes possible a self-controlled experiment. That is, some exits not used for sign modification can be instrumented to permit measurements before, during, and after the actual sign testing phase. Such measurements will indicate the degree to which the experiment may be biased by external conditions that are beyond experimental control or cognizance. They also supply the needed quantative information to correct for the bias to some degree. The use of the experimental site itself as a control group is desirable since certain characteristics of the actual population under test are evaluated over a long period to provide an indication of measurement stability. While it is conceivable that the signing experiment itself will affect the driving characteristics of the beltway traffic, such biases are expected to be smaller than those that might arise from the use of a totally distinct control group. For instance, local strikes, increases in the liquor tax, or other conditions imposed at the experiment site but not at the control site could result in a substantial experimental bias.

Because the use of highway signing is a direct component change countermeasure (Reference 7-15), it may be possible to evaluate the effectiveness of the sign modifications in terms of ultimate criteria (accidents, etc.). Thus, the use of accident data will occupy an important position in the countermeasure evaluation effort. The Highway Regulation Program Category, of all six program categories, is most susceptible to measurement in terms of ultimate criteria. However, as we have pointed out in Volume 2, the ultimate criteria do not adequately reflect the "why" of countermeasure operation; they only show the final result measured by a particular, rather insensitive standard. If the signing is to be effectively applied to other situations there must be other measures of effectiveness that detail the effect in measureable terms that can be translated into practical application guidelines. For the prototype experiment outlined here, two classes of measures are suggested:

1) Modification of traffic characteristics;
2) Driver reaction.

The purpose of measuring traffic characteristics is to determine the effect of the countermeasure in modifying those aspects of the flow that are deemed to be potentially dangerous (rapid braking, lane changing, large fluctuations in speed). In the experiment design it is assumed that such actions at an expressway exit may be caused by driver indecision as to what his proper course of action should be.

The use of driver reaction measures is a more indirect method of obtaining information on how the sign affects the driver. Here, surveys and questionnaires are employed in an effort to determine just how the signing affects the driver, and what he does in response to the information.

7.3.3 Site Selection

Important site selection considerations relating to the capability of local jurisdictions to administrate an evaluation program have been discussed in Section 9, Volume 2 of this report. Only those factors that deal with physical characteristics will be discussed here. Physical factors are easy to determine, but administrative considerations are probably more important for a successful experiment.

The prime requirement is to find a site where a recognized, definable problem exists. In the context of a highway regulation signing program this means finding a beltway with a sufficiently high traffic flow that has an exiting problem (accidents, complaints, known motorist confusion) amenable to solution by better signing (in the opinion of qualified highway engineering personnel). Although the determination of exits with high accident probabilities is one way of locating trouble spots, care should be taken to determine that the accident pattern is likely to be a result of the existing signing procedures. It is also necessary to insure that the location has a consistent period of problem behavior and that the site is not chosen simply on the basis of its recent (poor) accident history. Sites with very high accident peaks in their recent history are likely to improve without any experimental intervention because the one bad sample is not characteristic of the long term behavior.

In order to satisfy the requirements for statistically sound data, traffic density on the route must be quite high. In general this means a moderate-sized metropolitan area--a requirement that is consistent with the use of a beltway. By performing measurements at a number of exits, the required accumulation of accident cases could be obtained in a reasonable period of time. A freeway interchange in Ann Arbor, Michigan has been observed to have 40 accidents per year. If this behavior can be regarded as typical, a 20% reduction in accident rate on a 12 intersection beltway could be detected with some assurance in a period of one year.

The first step in selecting appropriate sites is to perform a survey to determine acceptable locations on the basis of size and existing problems. This survey should be performed by a team of highway engineers familiar with the problems of signing. Included in this survey should be a preliminary political contact with the local government agency having jurisdiction to determine if the operation of an evaluation program would be acceptable at that locality. From the list of basic possibilities produced by the evaluation team, the final selection would be made by competition; inviting those areas to submit designs for an experimental program using the guidelines furnished. This would permit each jurisdiction to propose a plan based on the innovative application of principles to the individual site while still retaining the basic desired plan outlined in the guidelines.

7.3.4 Selected Countermeasures

Basic countermeasures for the program are, of course, legend-bearing and symbolic signs recommended by the AASHO. The revised 1970 Manual of Uniform Traffic Control Devices has introduced a number of symbolic sign standards. In most cases, the conventional legend signs are an approved alternative--at least for the near future. Thus, three stages of signing can be employed:

1) Existing Signs (control);
2) AASHO Standard legend signs;
3) AASHO Standard symbolic signs.

Signing effectiveness is a function not only of the size and color of the signs, but also of placement and number. All of these factors will depend on the local situation and it is clear that all signs on a given beltway will not be uniform. Special situations such as left exits or dropped lanes might require special treatment. At any rate, the designs submitted by proposal can be analyzed by experienced traffic engineering personnel to determine the soundness of design in light of current knowledge.

7.3.5 Measures of Effectiveness

In Volume 13 of the Highway Safety Program Manual, the assertion is made that measuring the reduction in accidents is the most direct way to determine the effect of a highway safety program. While the use of ultimate criteria is possible for highway system changes where the countermeasure is closely related to the accident process, accident rates are usuallv insensitive to the "why" of countermeasure operation. Thus the use of intermediate criteria is necessary for understanding the immediate effects of the countermeasure and for extrapolating that effect to other situations with different physical and sociological characteristics.

To determine accident reduction, accidents must be measured on a before and after basis. Thus, about a year before the signing change is made, the recording of accidents at each instrumented exit on the beltway must be initiated. The exact period of time necessary to perform this accident determination will depend upon the accident rate at the experimental site and must be determined from available data once the site is chosen. This portion of the measurement phase is necessary so that natural fluctuations in the accident rate may be measured. Such information is important in determining if a change has actually occurred when the signing is introduced.

Two modifications to the standard accident data gathering procedure are suggested for the experiment: First, accident investigations should place special emphasis on the determination of signing as a factor in accident causation. In particular, additional variables should be added to the standard investigation form to permit the gathering and analysis of pertinent causative factors. An investigation team fashioned along the lines of the MDAI teams (but less sophisticated) might be useful in this respect. Second, an exposure survey to determine the characteristics of the driving population utilizing the beltway should be undertaken. The measurement of traffic conflicts (discussed below) is a type of exposure measure by some definitions

(Reference 7-16); however, a variety of other factors would be helpful in analysing the accident data. For instance, the determination of total mileage driven on the beltway as a function of driver age, sex, type of car, and other factors would allow a comparison of the population at fault in signing-induced accidents with the population at risk on the beltway. Only in this way can the relative involvement of different segments of the population be determined. The technique for making exposure surveys has been developed in a recent NHTSA study (Reference 7-17) and will not be repeated here.

As discussed earlier in Section 7.3.2, two types of intermediate criteria are suggested in addition to the measurement of accident reduction. These measures are intended to pinpoint the direct, anticipated effects of the countermeasure as well as any unanticipated effects that may occur. Thinking along the lines of a causal chain approach to accident occurrence, the most direct effect of a highway sign would be its perception by the driver and his understanding of the message to be transmitted. These factors determine his subsequent actions (and consequently the performance of the vehicle). Unfortunately such determinations cannot be performed on the large majority of highway travellers and only a small sampling could be obtained through the use preselected drivers.

The most direct, measureable approach to gauging the effectiveness of signs, therefore, seems to be a determination of what actions the vehicle undergoes in response to the drivers' understanding of the situation. That is, we expect the drivers' confusion over what exit to take (and other disturbing questions) to be reflected in the traffic characteristics. If a motorist feels he is passing the exit that he wants, he may brake suddenly and change lanes in an attempt to reach the exit; if he is unsure about the exit, he may slow down momentarily in an attempt to gain time to gather further information. Several techniques have been developed in the past to measure those characteristics of the traffic pattern that are potentially dangerous. The traffic conflict technique (Reference 7-8) has received considerable attention because it is simple to perform and takes little time to obtain statistically significant results. With this approach, the occurrence of "traffic conflicts" (as defined by actions such as braking or lane changing) is used to quantify the indecision felt by the driver when he passes the decision point. Another technique is the "acceleration noise" measurement (Reference 7-11) that is aimed at determining driver uncertainty by measuring the number of speed changes that occur at a given point per unit of time. The conjecture here is that a confused driver will slow down in an attempt to gain time to discern his proper behavior, then speed up again once the decision point is passed. This method has the disadvantage that it does not sample the driving population directly nor is it as inexpensive: Special instrumented cars are required to "float" with the traffic stream to provide the data. Some other techniques have been used (eg., galvanic skin response, see Reference 7-13) but they have not received the validation or the testing enjoyed by the two mentioned earlier.

The traffic conflict measurements need to be supplemented by more conventional measures of traffic such as density, flow rate, etc., in order to properly define the background conditions. That is, the number of conflicts can be expected to depend to some degree on traffic conditions: Braking conflicts are less likely to occur when traffic is light. In summary then, we suggest that the traffic modifications introduced by signing be determined by traffic conflict measurements with some attention to more conventional traffic characteristics.

A second less direct approach to determining the effectiveness of signing through the use of intermediate criteria is a study of driver reaction. The determination of driver reaction to signing is difficult to perform directly in a freeway environment as discussed above. If the experiment is conducted at an exit, then the usual procedure of stopping motorists for an on the spot interview may be employed to advantage. The disadvantage of this procedure is that the persons who have had the most trouble with the signing at this particular exit, may have passed the exit and thus are not available for questioning. The sample may thus be biased by people who have, by whatever maneuver, managed to make the exit.

While measurements of driver reaction in traffic are difficult to perform, there are a number of ways that subsidiary information may be gathered. One is the use of simulation experiments to expose a controlled sample of drivers to a photographed signing experiment in a realistic context (see References 7-7 and 7-14 for examples). While this exposure is certainly not equivalent to the tension-filled circumstances found in actual driving conditions, it does present the signs for comparative reaction by a sample of the driving population. Similarly, photographic simulation can be employed to determine reaction to color, size, and other physical factors as well as to determine psychological features such as glance readibility, comprehension and retention.

Although it is not strictly a measure of countermeasure effectiveness, it seems important in evaluation experiments of this type to take note of any external factors that may affect the measured results. For example, such economic and political factors as business recessions, riots or other civil disturbances, increased taxes, or labor strikes may affect people's behavior to such a degree that the results of the data may be in doubt. While it doesn't seem likely that factors of this nature can be accurately included in any analysis, a knowledge of their occurrence is important to determine the significance of the final conclusions.

7.3.6 Data Collection

In this section, specific techniques needed to implement the measures of effectiveness discussed above will be presented.

A modification of the normal accident reporting procedure to emphasize signing causation could be most conveniently performed by a supplementary form to be used in conjunction with the normal accident report. As discussed earlier, signing affects the driver directly and results in some response by the automobile. Any physical evidence to determine the vehicle trajectory during the accident process is consequently important in studying driver response. Important accident features include:

1) Skid marks on the pavement
2) Evidence of median or shoulder traversal (dirt, flattened grass, debris, etc.)
3) Damaged roadside features
4) Gore penetration

Much of this information could be collected most economically by photographs taken as soon as possible after the accident occurred, since vehicle disposition and certain other features of the accident (debris, etc.) may be modified quickly during the post-crash emergency operations. The aim of this phase of the accident investigation is to gather as much data as possible relative to the problem of determining the course of action taken by the driver in the accident process. This will help in verifying or rejecting accident histories obtained from more subjective sources.

The second portion of the investigation is concerned with personal interviews of the drivers involved in the accident as well as any other observers who were on hand at the accident. Care must be taken here to gather evidence that signing is a confusing or causative factor in the accident without specifically implying or asking for opinions on causation. Thus "Did you understand this exit (geometerics, connecting routes, etc.)"?, is a more appropriate question than "Did the highway signing contribute to causing this accident?". Some important questions to be determined from the drivers are:

1) Was the driver confused by the signing? In what way?
2) Did the driver see the appropriate signs?
3) Could he read all the information?

4) Could he understand the message?
   a) Were the required movements clear?
   b) Were the intersecting roads apparent?
   c) Was the point of application understood?
5) Did the driver follow the instructions? That is, did he believe them?
6) What happened at the accident point to cause the driver to leave his planned path?

Observers can similarly be questioned to yield as comprehensive a picture of the events leading up to the accident as possible (again without implying causation to bias their answers).

Data collection techniques for the suggested exposure survey are discussed in Reference 7-17. For the conflict technique, a minimum of four different pieces of information must be gathered: The traffic volume, the number of conflicts observed, the type of conflict, and the type of vehicle. A good example of the data gathering technique is presented in Reference 7-4. Conflict measurements are usually performed by direct observation of the intersection or by viewing photographic or television recordings. Sampling is performed at regular intervals throughout the day for periods that range from 15 minutes to one hour depending upon local conditions. Before beginning the actual data recording it is usually necessary to define the types of "conflict" to be categorized at the intersection in question since the particular geometry of the site may affect the type of conflicts that are obtained.

The number of measurements necessary to determine if a statistical difference in conflict rates (i.e. conflicts per vehicle) has occurred as a result of the signing change may be derived from the data taken during the control phase of measurements. As an example, the data on two-axle vehicles in Table I of Reference 7-4 can be used to illustrate the concepts involved.

In that experiment a set of conflict measurements taken with unmodified signing produced an average conflict rate of 0.042 conflicts per vehicle per hour with a standard deviation of 0.025. If the accident distribution is assumed to be normal, then Figure 7-1 shows the number of conflict measurements necessary to produce an indicated percentage change in the mean. Note that the percentage figures used in Figure 7-1 correspond only to the specific example used here and pertain to a 95% confidence level. Thus a total of 75 observations would be necessary in this case to detect a change of 20% (i.e. from 0.042 to 0.034 or 0.050) in the mean conflict rate for 95% of the experiments performed. Consequently, if the experiment is intended to detect changes in the mean on the order of 10%, hundreds of measurements would be required.

It is essential that such data requirement calculations be performed for both accident and conflict measurements to determine the number of samples (and hence the time required) necessary to obtain meaningful data. For the accident measurement, this information can be obtained from existing accidents records for the experimental site. For the conflict measurements the information must be estimated from a-priori knowledge of the site, or better derived from the data taken during the control phase of the experiment.

For the simulation phase of driver reaction to the signing, a number of tests might be performed. To determine the relative efficacy of legend or symbol signs to transmit different information types, the glance readibility of photographed signs can be compared for different viewing times. The simulations should be designed to correspond as closely as possible with the actual conditions involved in the experiment. Thus the messages used in the actual experiment can be displayed for times that correspond to real available viewing times. These tests should be performed with a sufficiently large set of subjects to determine if there is a significant difference in the perception of information for the different messages utilized. A simulation approaching actual conditions could be performed by photographing the actual signs in place on the highway from the driver's point of view. This film can be used in conjunction with a

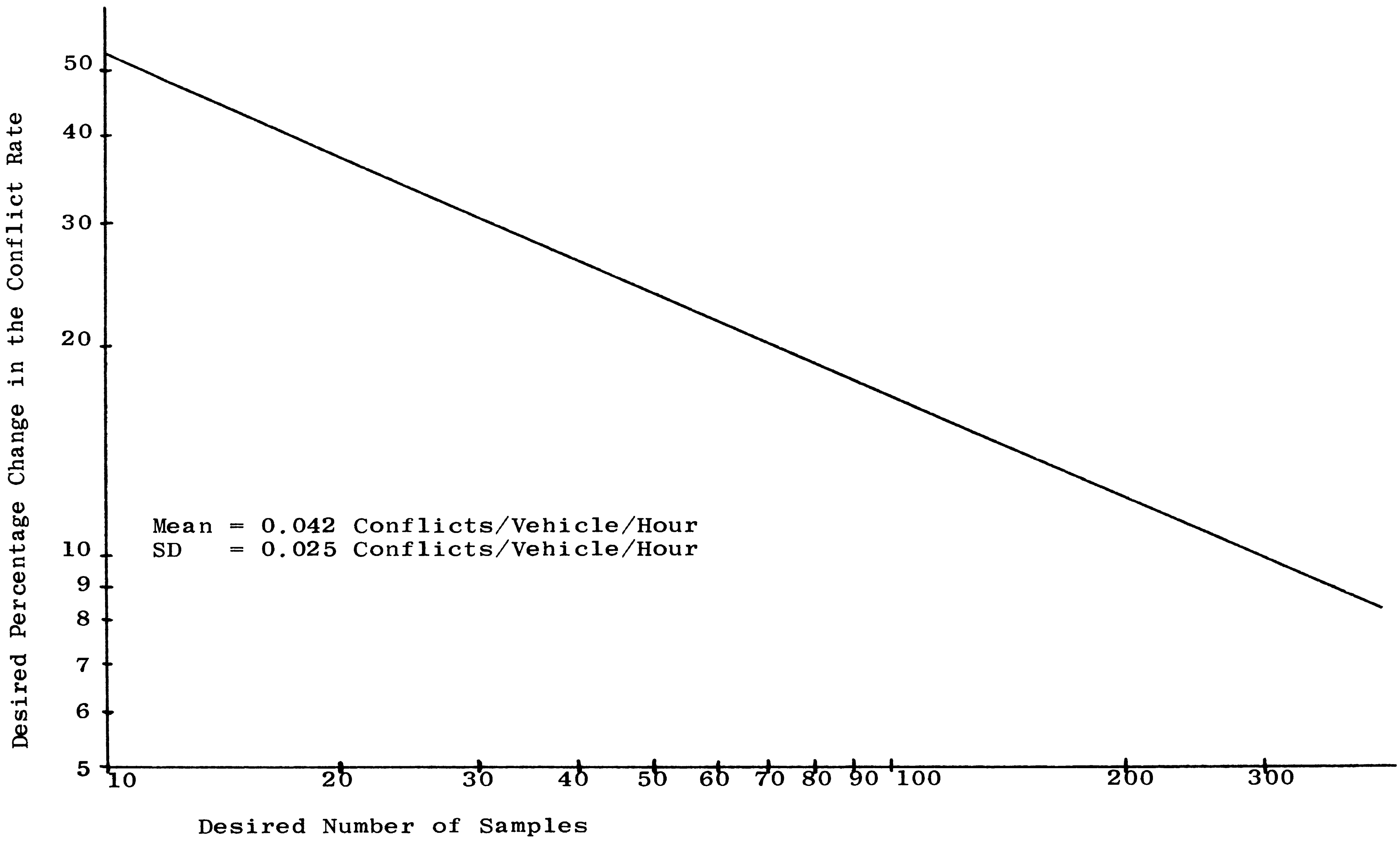


Figure 7-1 The Number of Samples Necessary to Obtain a Specific Change in Mean Conflict Rate.

questionnaire to determine the subject's response to the indicated instructions.

7.3.7 Data Analysis

Two important questions to be answered by the experimental program are:

1) Did the signing produce a statistically significant decrease in accidents? and
2) Did the signing produce a statistically significant decrease in conflicts?

These questions may be answered through the utilization of well-known statistical significance tests if the sample sizes were chosen sufficiently large to contract the effects of natural variability. Because of the close relationship between highway changes and the accident process, the measured reduction in accidents and the cost of the signing may be used to determine a cost-benefit figure for the experimental technique in question (conventional signs - symbolic signs). In the spirit of this study, however, blind faith in such figures must not rule out the possibility of other experimental effects that may be determined by the other data taking activities of the program.

One of the stated program goals is the determination of the relative effect of legend and symbolic signing. That is, are there some situations where legends are better and some where symbols are better? If so, what are they? This is connected with the <u>WHY</u> of countermeasure operation. That is, why did the signing produce a decrease (or increase) in the number of conflicts. Was it because drivers understood the new signing better and were able to make more careful decisions; or was it because they were so confused by the message they rejected it completely and were hence not confused at all? This determination of <u>WHY</u> must use all the available data (conflicts, accidents, interviews, etc.) to discern the real operating function of the signing. Only in this way can the principles be accurately applied to other conditions.

Another important question is the relationship between conflict rate and accident rate. If the conflict technique is to provide an accurate appraisal of accident potential, the relationship between conflicts, accidents, and other internal conditions must be determined.

Finally, it should be pointed out that if care is taken in the data collection process to carefully collect and record all available information in an accessible format (computer accessible magnetic tape, etc.), then the amount of data processing that can be performed is vast. It is important to note, however, that the recorded data must be accurate and thorough since analysis cannot reveal facts that are not there.

REFERENCES

7-1 A.Kuprijanow, S.Rosenweig, M.A.Warskow
Motorist Needs and Services on Interstate Highways
NCHRP Report Number 64, 1969

7-2 D.W.Schoppert, K.Moskowitz, S.F.Hulburt, A.Burg
Some Principles of Freeway Directional Signing Based on Motorists' Experiences
Highway Research Board Bulletin Number 244, 1960

7-3 S.G.Petersen, D.W.Schoppert
Motorists' Reactions to Signing on a Beltway
Highway Research Record Number 1970, 1967

7-4 A.W.Roberts
Diagrammatic Sign Study- Phase 1 Report
Report Number 70-015-7765
New Jersey Department of Transportation
May 29, 1970

7-5 Changeable Sign Messages
Rural and Urban Roads
Vol. 9, Number 1, January 1971

7-6 G.Johansson, F.Backlund
Drivers and Road Signs
Ergonomics
Vol. 13, Number 6, November 1970

7-7 W.J. Tierney, L.E.King
Traffic Signing - Symbols versus Words
A paper presented at the VI World Highway Conference
Montreal, Canada, October 4-10, 1970

7-8 S.R.Perkins
GMR Traffic Conflicts Technique Procedures Manual
Research Publication GMR-895
General Motors Corp.
August 11, 1969

7-9 R.E.Campbell, L.E.King
The Traffic Conflicts Technique Applied to Rural Intersections
Accident Analysis and Prevention
Vol. 2, pp.209-221, 1970

7-10 J.Pahl
A Comparison of Direct and Indirect Methods for Determining Accident Potential
Accident Analysis and Prevention
Vol. 2, pp.201-207, 1970

7-11 C.L.Heimbach, H.D.Vick
Relating Change of Highway Speed per Unit of Time to Motor Vehicle Accident Rates
Highway Research Record Number 225, 1968

7-12 J.F.Torres
Acceleration Noise, Power Spectra, and Other Statistics Derived from Instrumented Vehicle Measurements Under Freeway Driving Conditions
Report Number TM-3858/025/00
System Development Corporation
September 30, 1969

7-13 R.M.Michaels
Effect of Expressway Design on Driver Tension Responses
Highway Research Board Bulletin Number 330, 1962

7-14 A.Burg, S.F.Hulbert
Predicting the Effectiveness of Highway Signs
Highway Research Board Bulletin Number 324

7-15 W.Carlson, W.Hall
HIGHWAY Safety Project Evaluation Methodologies
Highway Safety Research Institute, The University of Michigan
December 31, 1968

7-16 P.S.Carroll
The Meaning of Driving Exposure
HIT LAB Reports, April 1971
Highway Safety Research Institute, The University of Michigan

7-17 P.S. Carroll, T.L.McDole, W.L.Carlson, P.F.Samarco
Acquisition of Information on Exposure and on Non-Fatal Crashes
Volume III-Procedures for an Exposure Survey
Highway Safety Research Institute, The University of Michigan
May 12, 1971
DOT Contract FH-11-7293

APPENDIX A

AUTO LIGHTING SURVEY PROCEDURES

1. OBJECTIVES. The purpose of the survey is to determine the proportion of automobiles in the different states which have certain readily observable defects. The larger purpose of the study is to develop one measure of the effectiveness of various periodic Motor Vehicle Inspection Programs.

2. LOCATION. There are a number of possible sites for taking the survey: expressway exits, stop-sign intersections, plant gates, and shopping center exits. The latter is the preferred location, but alternatives should be considered in consultation with the survey director. Using the traffic counter described under equipment, the location should have a single lane in the direction of the surveyed cars and preferably a raised lane divider. With two-man crews using a hand counter, this requirement may be dropped, but only in the case that no alternative can be found since weaving movements in multilane situations may complicate the data gathering task. Once the general site has been selected, the exact spot for the survey should be chosen on the basis of ease of observation of cars, possible inconspicuous placement of observers, moderate traffic volume (e.g. a side rather than a main exit of a shopping center) and observer's personal safety from traffic hazards. In all cases permission to conduct the survey should be obtained from the appropriate person: local police or traffic authority (public streets), plant manager (plant sites) or center manager (shopping center). The center manager may usually be reached by contacting one of the stores in the center and determining his name and address. In approaching the proper authority, the surveyor should identify himself and give a brief explanation of the purposes of the survey. In all cases it should be stressed that the operation will involve no obstruction of the flow of traffic and, in the

case of private property, will not involve any police action. In the case of public streets, one should obtain promise of non-interference with the survey by police in order not to bias the results; if the police are present and/or issuing tickets for observed defects this will bias the sample through a "check-lane avoidance" effect and also would tend to pose public relations problems. In all cases request the authorities not to give the survey an advance or after the fact publicity. In the unlikely event that the responsible person appears reluctant to cooperate, offer to provide him with a traffic count and results of the survey for his location; if he so desires, promise that the specific location will not be revealed. If he still appears reluctant one may use his judgment either to select an alternative site or to attempt to persuade further the responsible individual.

3. FORMS. The surveyor will be provided with two forms; a car population survey form and a defect survey form. He will receive one copy of the population form for each survey site and a pad of defect survey forms for the surveys. The forms are fairly simple and the surveyor should not have difficulty in following them. An explanation of abbreviations follows and instructions on entering data are given in sections 5 and 6. All locational data should be filled out on both forms. If taking more than one survey each page of the defect form should have appropriate identification information and the first page should have complete information. The "Area" designation of the defect form refers to the type surroundings in which the survey was conducted such as "commercial" "shopping center" "industrial" "residential." "Intersection" should give a complete characterization of the intersection used, i.e. "Western exit, single, right-turn-only exit lane into four-lane divided highway." In the defect column the abbreviation "ot" means unlit bulb which should be illuminated, "brk" means a broken, shattered, or missing light component, "muffler" means either a noisy muffler or obvious exhaust system defect such as fumes escaping from beneath the passenger compartment, "wiring" refers to other obvious

faults of the lighting system which appear attributable to wiring defects such as back-up lights on when driving forward, shorting flickering lights, parking lights burning while headlights are on, etc. "Glass" refers to any broken, spiderwebbed, or grossly obscured front, side, or back glass.

4. EQUIPMENT. The equipment to be used will be demonstrated by an experienced user to new personnel. The equipment will include for a single-man survey an adequate supply of forms, a copy of these procedures, a clip board, a flashlight, the traffic counter, and if conditions warrant a voice recorder. For the two-man crew, the basic equipment will remain the same except a hand counter may be substituted for the traffic counter. All equipment will be placed in a carrying case in kit form.

5. POPULATION SURVEY. The first survey to be conducted is that of the auto population from which the defect study will be approximately drawn. The auto population survey data sheet contains a matrix of common auto makes and years. For domestic manufacturers, most major makes are listed. If a particular sub-make is not listed, e.g. Chevelle, it should be listed under the appropriate major make, Chevrolet. If one cannot find an appropriate category for a given domestic make i.e. Henry J, or Allstate, place a tally in the Special Vehicles category and circle it. For out-of-state cars enter a tally in the column marked "out" but not under the year column. All foreign cars, trucks, and special vehicles should be listed in the appropriate row, but the model year should be disregarded; simply enter a tally for each of such items starting in the '67 column. If there are any questions of classification indicate a description of the car on the reverse of the form and call it to the attention of the survey director. In parking lots, sample the first five cars on each side of each aisle in the proximity of the exit from which one will be conducting the defect survey. If parking configurations do not allow this exact pattern an alternative survey pattern may be devised. In any case at least

100 cars should be included in the population survey. In street locations, the surveyor should before beginning the defect survey sample the cars passing on the basis of 1 of 3 or 1 of 5 to get an adequate measure of the car population. For shopping center car population survey one might alternatively sample the cars entering the parking lot on the same basis as the street population sample.

6. DEFECT SURVEY. Once it has become sufficiently dark for most vehicles to be displaying their headlights, the surveyor should begin the defect count. In any case of a single-man survey or a two-man survey using the traffic counters, the counter should be set up in the following manner: the light source should be placed on the left hand side of the traffic lane aimed at the photo cell counter mechanism. The light source should be placed on the lane divider. If no lane divider exists, the surveyor should obtain a construction type flasher baracade to protect the light source, but should primarily seek a lane provided with a lane divider. The photo cell-counter unit should be placed on the curb side and, if possible, approximately 10 feet from the curb. The entire unit should be placed sufficiently far away from the intersection to prevent its being triggered by lights of turning and intersecting cars. The counter should be adjusted so that it is aimed at door panel height. A good rule of thumb is a knee height above the roadway. Surveying should begin immediately after the counting mechanism has been activated. The surveyor should position himself from 25 to 40 feet away from the stopping point of the intersection and far enough from the curb to be inconspicuous yet still have a full view of the approach of the intersection. For uniformity several rules should be followed. All out-of-state cars should be counted, but no entry should be made for them except a marginal notation to tally their total number to correct the count. When motor cycles approach the counter, the surveyor should note the number of counts which the motor cycle makes since motor cycles are not to be included in the survey, and they may register multiple counts on

the traffic counter. Note the make and year of the defective vehicles to the best of your ability. Do not concern yourself with the driver age and sex characteristics. If anything is wrong with any item in any set of lights mark it as a single defect. Example: four headlight car with both left headlights burning and both right headlights out would count as a single defect or a car with three tail lights on each side with one of the left tail lights out would be a single defect. If the car has not yet turned on its headlights, the benefit of the doubt should be given and one should check for brake light defects only. If both tail or both brake lights are out the appropriate row for this condition should be checked rather than indicating in the separate rows for each side. If a combination brake and tail light is broken enter in the appropriate broken tail light row. If both the brake and tail light parts of a combination fail to light indicate each defect in the appropriate row. If defects are noted for a car and then it is observed to be an out-of-state car place a large X mark in the "outstate" row, and also include the car in the out-of-state tally for counter correction. All of the foregoing instructions on data gathering apply to all techniques used, form, voice recorder, single-man and two-man crews. In the case of two-man crews without the traffic counter one man should record defects and the other should count cars with a hand counter. The voice recorder should be used as a supplementary device when traffic flow is heavy. The surveyor should have a form in sight to follow the general format. It is suggested that as many cars as possible be recorded on the form. However, when it is necessary to use the voice recorder all entries for the period during which it is used should be with it. Care should be taken to avoid double counting and over one period of time one should not try to mix recording medias. Do not try to keep a continuous written record using the recorder as a catch up device. After the survey period is over, the surveyor should note on the first page the total number of cars counted and the number of out-of-state and other skipped vehicles, e.g. motor cycles. The traffic counter should be disassembled and returned to the case.

7. GENERAL. It is advised that the surveyor wear casual clothes to avoid an official appearance which might produce a "check-lane avoidance" effect. He should have his observation point in a well lighted area. If car occupants ask him what he is doing reply "Highway Safety Survey." If passers-by make inquiries and are not satisfied with the answer given to car occupants wait until a lull and give them an outline of the project as stated in the objectives of the survey. Do not engage in protracted conversations and do not debate relative merits of various safety programs. If they are still curious give them a card and ask them to write. If conditions become unsafe at the observation point (thunderstorm, hostile driver response, attempted theft, etc.) do not hesitate to make a strategic withdrawl leaving equipment if necessary since the surveyor's personal well being and safety and the need of avoidnng incidents exceeds the value of equipment and any data which might have been collected.

If one is surveying the same shopping center more than one night in a row one should discontinue the survey at store closing time on the second and subsequent nights since inclusion of cars after that time on the later nights will bias the sample by counting employees' cars more than once.

APPENDIX B

SAMPLE QUESTIONNAIRE OF VEHICLE OWNERS FOR VEHICLE REGULATION

Instructions to Respondents:

This survey is being conducted in your city in an attempt to learn more about how people feel about highway safety. We ask that you complete it fully and thoughtfully. If you feel uncertain about the answer to some question, give the answer that you feel best represents your ideas. All answers will be kept confidential.

Name:________________________________ Age:_______________

Adress:______________________________ Sex:_______________

Occupation:__________________________(Clerk, teacher, etc.)

Education:___________________________ (High school, college, etc.)

We want to know something about your driving habits and about the vehicles which are used by your household on a regular basis. Please list the information on every vehicle that you or members of your family living with you here own, lease, or have principal control of. List the car which you most often drive first.

Make	Model	Type	Year	Ownership	Current Mileage
________	________	________	________	________	________
________	________	________	________	________	________
________	________	________	________	________	________

List additional cars on the back of this form.

We want to know something about how often you maintain the vehicle which you listed first. Circle the correct answer.

Who takes responsibility for servicing this vehicle on a regular basis? Who, in other words, decides where, when, and how much service is performed?

1. You
2. Another family member
3. Leasing company (for leased car)
4. Employer (for company car)

When was the last time you had this car repaired?

1. This week
2. Sometime in the past month
3. Sometime in the past six months
4. Sometime in the past year
5. Sometime before the past year
6. No opinion

When was the last time you had this car regularly serviced and checked? That is when did you last have an oil change, lubrication, tune-up, etc.?

1. This week
2. Sometime in the past month
3. Sometime in the past six months
4. Sometime in the past year
5. Sometime before the past year
6. No opinion

Please indicate which of the following items were repaired the last time that the car was repaired. If more than one item was repaired, circle all items that apply.

1. Brakes
2. Steering
3. Tires
4. Lights
5. Muffler
6. Shocks or springs
7. Engine
8. Transmission
9. Other
10. No opinion

How often do you think that your car should be checked and serviced?

1. Once a month
2. Once every three months
3. Once every six months
4. Once a year
5. When needed
6. No opinion

How often do you have your car serviced and checked?

1. Once a month
2. Once every three months
3. Once every six months
4. Once a year
5. When needed
6. No opinion

Please indicate what is your best opinion about the condition of the following items on your car. Check for each item the column which is your belief.

Item	Excellent	Good	Fair	Poor	Very Bad	Don't Know
Brakes	___	___	___	___	___	___
Tires	___	___	___	___	___	___
Steering	___	___	___	___	___	___
Shocks and Springs	___	___	___	___	___	___
Lights	___	___	___	___	___	___
Wipers and Washers	___	___	___	___	___	___
Muffler and Exhaust	___	___	___	___	___	___
Engine	___	___	___	___	___	___
Transmission	___	___	___	___	___	___
Body	___	___	___	___	___	___

Many people work on their own cars to make repairs. Would you say that you usually

1. Never make repairs?
2. Make minor repairs?
3. Make both minor and major repairs except for some special items like transmission or body work?
4. Make all repairs?
5. No opinion

Have you ever had training or experience as an automobile mechanic?

1. Neither training nor experience
2. Training only
3. Experience only
4. Both training and experience

Do you presently work as an automobile mechanic or other automobile service personnel?

1. Yes
2. No

Do you work on your car as a hobby?

1. Yes
2. No

We would like to know your opinions about several things connected with cars. We have listed a number of statements, and we want to know how much you agree with them. Please check the appropriate column by the statement.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Most people I know take very good care of their cars.	___	___	___	___	___
Cars in bad condition cause a lot of accidents.	___	___	___	___	___
When I drive, I worry a lot about accidents.	___	___	___	___	___
I like bumper stickers, they let people know where you stand.	___	___	___	___	___
My car is in pretty good condition.	___	___	___	___	___
You can get by with slick tires, if you stay off the expressways	___	___	___	___	___
Most people I know check their cars regularly for mechanical defects.	___	___	___	___	___

	Strongly Agree	Agree	Neu-tral	Dis-agree	Strongly Disagree
No matter what you do, sooner or later you'll have an accident.	___	___	___	___	___
Most accidents are caused by careless drivers.	___	___	___	___	___
What they really need to do to stop accidents is to get junkers off the road.	___	___	___	___	___
I check my brake lights every two weeks.	___	___	___	___	___
The police around here are really tough on bad cars.	___	___	___	___	___
Most of the time you can let things go on a car without any danger.	___	___	___	___	___
When something is wrong with their car, most people I know adjust their driving to take account of it.	___	___	___	___	___
In general, the garage where I go does a good job.	___	___	___	___	___
Speed kills.	___	___	___	___	___

The following set of questions lists some common problems that most automobile owners have experienced from time to time. We would like to know what you think has caused these problems.

Steering wheel vibrates or shimmies at high speeds.

1. Soft rear tires
2. Loose clutch plate
3. Misaligned wheels
4. Broken universal joint
5. No opinion

Car continues to bounce around after hitting a dip.

1. Uneven load distribution
2. Unbalanced wheels
3. Engine out of tune
4. Worn shock absorbers
5. No opinion

Turn signal indicator burns steadily rather than flashing.

1. Broken universal joint
2. Blown electrical fuse
3. Burned out brake light
4. Loose distributor cap
5. No opinion

Car pulls to one side when stopping.

1. Worn or dirty brake shoes
2. Overinflated tires
3. Loose king pin
4. Worn shock absorbers
5. No opinion

Tires wear out more at the sides than in the center.

1. Misaligned frame
2. Under inflated tires
3. Over inflated tires
4. Unbalanced wheels
5. No opinion

Brakes have to be pumped to make car stop.

1. Broken universal joint
2. Possibly two or more flat tires
3. Ruptured brake lines
4. Not enough brake fluid
5. No opinion

Lights flicker at night.

1. Faulty dimmer switch
2. Loose fan belt
3. Faulty voltage regulator
4. Blown fuse
5. No opinion

One front wheel appears to wobble.

1. Defective ball joints
2. Unbalanced tires
3. Loose pinion arm
4. Broken universal joint
5. No opinion

Car won't move even though engine is running and car is in gear.

1. Loose pinion arm
2. Broken universal joint
3. Overinflated tires
4. Loose fan belt
5. No opinion

Loud buzzing sound is heard when door is opened.

1. Turn signals left on
2. Blown fuse
3. Loose fan belt
4. Keys left in ignition

Tires have a zig-zag pattern of tread wear.

1. Blown fuse
2. Broken universal joint
3. Loose idler arms
4. Over inflated tires
5. No opinion

Both taillights or both headlights go out at once.

1. Dead battery
2. Faulty voltage regulator
3. Blown fuse
4. Loose ball joints
5. No opinion

Indicate whether you think the following statements are true or false. If you think a question is mostly true or mostly false give the answer which you think is more correct.

	True	False
A loud muffler is a safety hazard	____	____
Letting the air out of tires gives you more traction	____	____
Soft tires run cooler	____	____
Rear brakes wear out sooner since they must pull the car to a stop	____	____
Unbalanced tires cause a thumping sound	____	____
Streaking windshield wipers are caused by incorrect arm tension	____	____
A slight rainbow discoloration at the edge of windshields is normal and can be ignored	____	____

	True	False
You can tell if your headlights are misaimed by looking at the bulb	___	___
The steering wheel should turn a couple of inches without moving the tires when the car is stopped and the engine is off	___	___

The following questions deal with the motor vehicle inspection system that is (will be) operating in your area. Some of the questions deal with factual knowledge about the system and some with your opinion of the system. In each case give your best judgment of the correct answer or the one that most fits your feelings.

I first learned of the system from

1. Newspapers
2. Radio or TV
3. Friends
4. Service station or garage
5. Police
6. Motor vehicle department
7. This interview
8. No opinion

I first learned of the system during

1. Today
2. This week
3. This month
4. This year
5. Sometime else
6. No opinion

To have my car inspected, I must

1. Do it myself
2. Take it to an authorized garage or service station
3. Take it to a government inspection station
4. Allow the police to check it at any time
5. Both ____ and ____ of the above
6. No opinion

I must have my car inspected

1. Once every 3 months
2. Once every 6 months
3. Once every year
4. When requested by police
5. At least once a year but more often if required to do so because of minor problems
6. Both ___ and ___ of the above
7. No opinion

A fair charge for the inspection is

1. Nothing
2. Less than $1.00
3. Between $1.00 and $5.00
4. Between $5.00 and $10.00
5. More than $10.00
6. No opinion

I think that people making inspections are

1. Fair and correct in their decisions
2. Trying to do a good job, but missing a lot
3. Out to take advantage of people
4. No opinion

The system that I would like best would be

1. No inspection
2. Self-inspection
3. A system run by service stations and garages under state supervision
4. Inspection at a government station
5. No opinion

I think that cars should be checked by inspection

1. Once every two years
2. Once every year
3. Once every six months
4. Once every three months
5. No inspection
6. No opinion

Most of my friends think that the system is

1. A real burden
2. Some bother
3. No inconvenience
4. Of some service
5. Best thing they could do to stop accidents

I think that the system will do a lot to save lives

1. Strongly agree
2. Agree
3. Neutral
4. Disagree
5. Strongly disagree

Most of my friends have had trouble getting their cars through inspection

1. Strongly agree
2. Agree
3. Neutral
4. Disagree
5. Strongly disagree

The system is just a way of the garages getting more money out of poor people

1. Strongly agree
2. Agree
3. Neutral
4. Disagree
5. Strongly disagree

A lot of people around here try to beat the system

1. Strongly agree
2. Agree
3. Neutral
4. Disagree
5. Strongly disagree

APPENDIX C

DETERMINATION OF NUMBER OF CASES

As a first approximation the serious emergency cases can be treated as a binomially distributed random variable--those in which errors occurred and those in which errors did not occur. If we knew the probability (p) of an attendant error, it would then be possible to define (mathematically) the probability function of any sample we might take. In order to gain some insight we have considered the effect on the sample distribution for populations having a p in the range 0.40 to 0.60. It is well known (Hogg & Craig*) that the sample drawn from a binomially distributed population (N*p) can be approximated by a normally distributed random variable, for large N. The mean is the proportion of the sample having the characteristic of interest (e.g., error in diagnosis or treatment). When this approximation holds it is possible to establish a confidence interval about the difference between two sample proportions such that the probability of the true difference lying within the interval is known. This confidence interval can be found from the following relationship:

$$C.I. = \hat{p}_1 - \hat{p}_2 \pm Z_\alpha \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{N_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{N_2}} \qquad (1)$$

*R.V. Hogg and A.T. Craig, Introduction to Mathematical Statistics, 3rd Edition, 1970, The MacMillan Co., Toronto, p. 201.

$\hat{p}_1$ - Observed proportion in the sample from population 1 (before the program).

$\hat{p}_2$ - Observed proportion in the sample from population 2 (after the program).

N_1 - Sample size for population 1.

N_2 - Sample size for population 2.

Z_α - Normalized deviation for a 1 - α confidence interval.

If the assumption is made that both samples contain an equal number of observations, a Minimum Significant Difference (M.S.D.) can be obtained from the following relationship

$$\Delta p^2 = Z_\alpha^2 \frac{\hat{p}_1(1 - \hat{p}_1) + (\hat{p}_1 - \Delta p)(1 - \hat{p}_1 + \Delta p)}{N} \qquad (2)$$

$N = N_1 = N_2$

$\Delta p = \hat{p}_1 - \hat{p}_2$

Δp - Minimum Significant Difference (M.S.D.)

By appropriate algebraic manipulations and the application of the Quadratic formula the MSD can be obtained from

$$\Delta p = \frac{\frac{Z_\alpha^2}{N + Z_\alpha^2} + \sqrt{\frac{Z_\alpha^2}{N + Z_\alpha^2} + \frac{8\hat{p}_1(1 - \hat{p}_1)Z_\alpha^2}{N + Z_\alpha^2}}}{2} \qquad (3)$$

Figure 6-4 indicates the relationship between sample size and the MSD. This figure assumes a one-tailed significance test (e.g., the hypothesis that there is no change in the proportion of cases with errors) is tested against the hypothesis that there is a reduction in the proportion of cases with errors. Based on these assumptions, a sample size of 300 will enable us to detect a change if the observed occurrence of **cases with errors** drops from p = 0.50 to 0.39. By the same assumptions a doubling of the survey from 300 to 600 observations would enable us to detect a change if the observed occurrence of treatment errors drops from p = 0.50 to 0.42. These particular values assume an initial p of 0.50. But as can be seen from Figure 6-4 the relative magnitudes of the MSD given other p's will be proportional. Thus it is concluded that the potential improvement resulting from a doubling of the survey would not be justified. Thus, the survey is designed to obtain approximately 300 observations.