16407

Report HSRI-71-117

PLANNING FOR CONCENTRATED IMPLEMENTATION OF HIGHWAY SAFETY COUNTERMEASURES

Volume 1: Introduction and Summary

J. O'Day

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. I		

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

15. Supplementary Notes

16. Abstract

Although the sixteen Highway Safety Program Standards have been in effect for some time, their adoption by the individual states has been less than complete. To convince the local governments that the requirements imposed by these Standards are well worth their cost in improved safety benefits, a need exists for determining and ranking the efficacy of highway safety countermeasures. The misconceptions, pitfalls, and potential success that abound in the evaluation of social system changes constitute the initial subject of this volume. It is argued that the careful, but extensive, measurement of all pertinent factors in the action of a highway safety countermeasure is a necessary approach to determining efficacy. Based upon this measurement-oriented approach, a summary of the projected program plans for the sixteen Standard areas is presented. Six program categories are defined which cover the range of activities defined by the Standards; moreover, these program categories group countermeasure types that are amenable to simultaneous evaluation. Program plans for each of the six evaluation experiments are described briefly in this volume. 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	36+ vi	

Form DOT F 1700.7 (8-69)

PREFACE

This document is the first 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 an introduction to the evaluation of countermeasures and a brief summary of results. Volume II contains the rationale for program planning; Volume III describes the detailed program plans for six countermeasure evaluation categories; and Volume IV is a report bibliography on documents pertinent to countermeasure development.

TABLE OF CONTENTS

FIGURES

TABLES

1. INTRODUCTION

There have been 16 Highway Safety Program Standards promulgated by the federal government, along with suggestions of particular activities which might be undertaken by local or state agencies to reduce the incidence of highway crashes. Although few people question either the spirit of the Standards, or the concept that the recommended activities are useful, many question the cost-value relationships. There is a clear need for information when some potential program organizer asks: "Will this activity be worth as much (or more) to me as it costs?"; or "Which of the many programs which could be undertaken will be the most useful?"

Since most of these activities (often called accident countermeasures by the people who promote them) must ultimately be instituted by local and/or state agencies, these agencies and their funding sources must be convinced of the expected value of a program before it begins. Thus it follows that most often the information proving the value of a program, or indicating its potential value, should be in a form most palatable to the decision-makers.

In early 1970 the staff of the National Highway Safety Bureau* (NHSB) developed a plan entitled "Safety through Concentrated Operational Program Effort"--subsequently referred to as SCOPE. Its purpose was to "implement operation evaluation programs and to support sustaining programs, in order to develop a priority scheme for use in the allocation of funds to various elements of a traffic safety program at the Federal, State, and local level" (quoted from an undated NHSB working document prepared early in 1970).

The state-of-the-art review conducted at that time led to the conclusion that there was simply a lack of valid data on the many variables associated with the traffic safety field, and that this lack constituted an impediment to the estimation of payoffs on selected programs. Measures of the effectiveness of traffic safety programs to that date were still too fragmented to form a basis for priority decisions at the national level.

The fact that this problem still exists, at least at the state level, is clearly indicated in a recent letter from a director of a State Office of Planning and Programming. The letter was in response to a rather general question about the problems of resource allocation within the highway safety field and indicates the dilemma still faced by such planners:

> The whole business of planning is a very difficult concept to sell or promote. People seem to have a built-in affinity to action. The crisis nature of our society has something to do with it. There are some problems in relation to highway safety planning that clearly need consideration. The commentary that follows describes a little of what we cope with daily.
>
> Evaluation techniques must be relatively simple and readily adaptable to the personnel levels at state and local levels. While we have must information on drivers, vehicles, etc., often it is not in a status to be manipulated, as is our case now. States have been great at collecting information as required by the codes but have done little in terms of management analysis. We readily assume that if 500 people go through a driver improve-

*In January 1971 NHSB became the National Highway Traffic Safety Administration (NHTSA).

ment program and only 50 have violations in the next 12 months, that 450 obviously benefitted. Few realize that the same thing could happen if we did nothing to those 500 drivers. Programs like driver education are infinitely difficult to evaluate. We believe Iowa has a top quality program in terms of content, coverage and volume. Yet, we cannot support it to the Governor and Legislature.

Take another issue. What is the effect of adding 100 highway patrolmen? What is the effect of installing an emergency telephone system on Iowa's Interstate System? What do we know about PMVI? We have noted some evidence that the accident, injury and fatality rates often go up after PMVI is initiated. I doubt if many states have base data from which to start evaluating PMVI. Iowa certainly doesn't. How do you determine the percentage of accidents deterred by inspection of a certain number of items on a selected population of vehicles? We have no base data or benchmarks to start from and with the partial inspection program legislated, such effects will be very difficult to statistically determine even after 5 years. Vehicle registration in most states has been a revenue producer, not a safety program. The information has helped criminal law enforcement.

Our Highway Commission is about to let a contract for the design of an accident locator system of more detail than the existing scheme. It will be some years before we have accident records of greater reliability than those in existence. We have had occasion to think about punching up the accident records of the last 10-12 years at the Iowa State Computer Center. However, having some experience in the reliability of questionnaire-type responses (and that is what these really are) we simply couldn't see the chance of reliable results at a cost of $30,000.

We have observed that often a program evaluator has aimed too high in his measurement of a program. For example, there is the obvious hope that a change in the driver education program will result in an immediate decrease in the state's fatality rate, but it is more likely that the effect of such a change might only cause a small change for the better over the lifetime of the individual driver. While the effect may be real it will be rather difficult to measure. Perhaps it would be better to aim somewhat lower, but with the possibility of drawing a firm conclusion regarding a program's results. It might be that two well-controlled groups of students could have their violation and accident records compared statistically leading to the conclusion that there was indeed an effective change.

But to be able to draw such conclusions validly requires considerable attention to detail in the design and conduct of the program. Experimentation within the social system is dangerous at best, because there are almost always extraneous factors contaminating the results. If one wished to compare two groups of driving students, the evaluation might be contaminated by the decision of one school system to modify the student parking regulations (i.e., so that more students drove to school in the "after" period). It is not certain that such interference cannot be tolerated, but such details must be taken into consideration before drawing conclusions.

We have suggested in this report that a systematic approach to the evaluation of highway accident countermeasures is much to be desired--in fact, that it is the best way to gather the evidence desired in the most useful form. This implies that there should be some plan or structure for the conduct of experimental countermeasures so that the information obtained is both that which is needed (to prove or disprove a point) and in the form in which it can best be used.

2. QUESTIONS AND ANSWERS ABOUT THE RECOMMENDED DEMONSTRATION PROGRAMS

As the introduction has been somewhat abstract, it should have generated in the reader's mind some questions which we have tried to anticipate. We will proceed with both some questions and their answers.

Q. What is a "systems (or systematic) approach" and how can it help to evaluate real highway safety programs?

A. The term "systems approach" is perhaps overused, but in its most general sense it simply means thinking about a problem abstractly before making recommendations for change. The thinking may involve a large amount of engineering or economic activity, and it may take a lot of time and money; or it may simply involve the considered judgement of a qualified expert. For simple problems the expert may appear to have arrived without effort at recommended changes in the present system; but you will find that he usually does consider problems in an abstract sense--although if he is familiar enough with the problem he may do so in his head.

The genuinely expert auto mechanic is a systems engineer; not so the pseudo-expert mechanic who is, in reality, a tinkerer. If you have taken your car to a mechanic because it sometimes does not start, the pseudo-expert may more or less randomly suggest a "cause" and perhaps sell you a set of spark plugs; if that does not seem to do the job he may try, at your expense, a new battery, an oil change, and a tune-up. After all these changes have been made, the car may or may not start more easily. The real expert, following the systems approach, may begin by asking you whether it starts poorly all of the time or just under certain conditions, say in wet weather. He will also make a careful examination of the car; he will note the oil on the spark plugs and wiring harness and will correctly advise you to have the engine steam cleaned and a new rocker arm cover gasket installed.

Most true experts in the field of auto mechanics or elsewhere will take the "systems approach" whether they call it that or not. But when a system becomes very large--like a country's telephone system, or the highway system--the prediction of the effect of a change calls for specialized techniques. These include modeling (setting up an abstraction of the problem which can be used to investigate "what if" sorts of questions on paper), simulation (the same sort of method using computers), laboratory tests (in which there is relatively good control of the experiment albeit under somewhat unrealistic conditions), and field tests or demonstrations (where the control is not as good as in a laboratory, but the realism is likely to be better).

This report is primarily concerned with the problem of measuring the effects of changes in the highway traffic system. In terms of the diagram in Figure 1, it is concerned with "tests and demonstrations" shown as part of the abstract or make-believe world--the results of which will hopefully lead us to a better real world.

Q. How will such tests and demonstrations help the highway accident problem?

A. A principal problem with regard to the many standards which have been proposed is that they still must be "sold" to the potential customers--the cities and states of the country who must implement the programs. And there are still a good many people who can be sold best if they can be shown that something really works. One way to sell a safety program is to <u>prove</u> that it is a good idea; and to prove it in terms that the decision-makers can understand and appreciate.

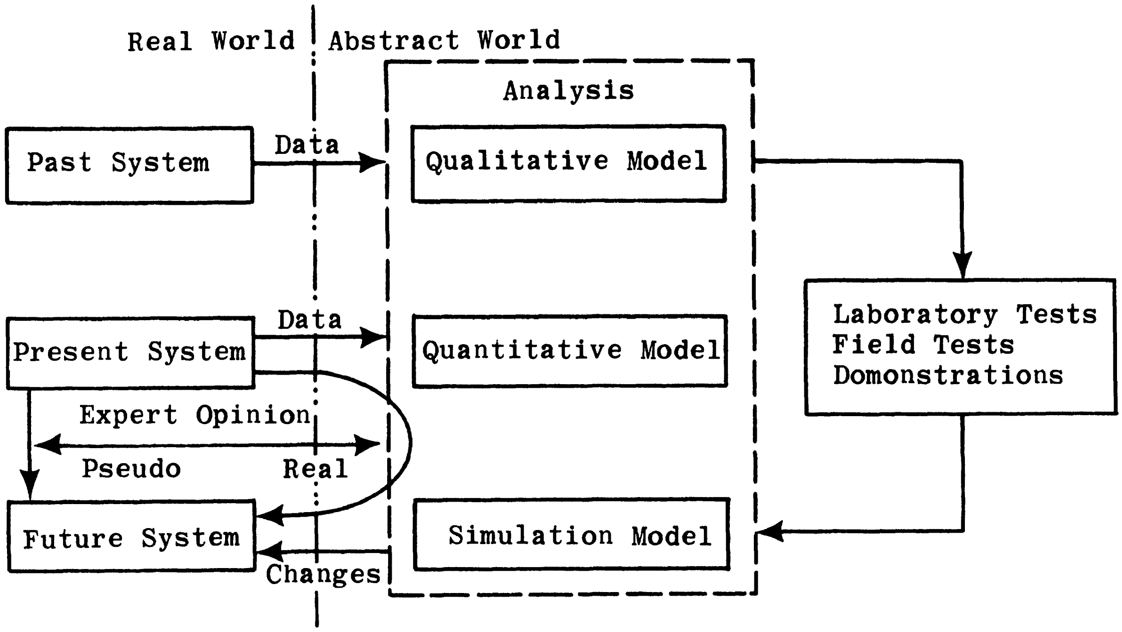


FIGURE 1. THE SYSTEM DESIGN PROCESS

Havelock and Markowitz (1), in a recent study for the National Highway Traffic Safety Administration (NHTSA), discussed at length the problem of improving communications between researchers and decision-makers; they were concerned with the problem of utilization of research findings. In questioning decision-makers they found the most frequent need expressed was for the translation of research findings into understandable terms. In their study, decision-makers were important people of varied but seldom scientific backgrounds, representing legislators, industry executives, public officials, etc. We suggest that demonstration programs have the potential for being particularly effective in persuading decision-makers to act, but only if the results of the demonstration can be put into an easily understood language. That is, it will be most useful if the results of an experimental program are supported by statistical considerations <u>and</u> are made quite clear to the reader unfamiliar with such concepts.
Q. Has anyone ever conducted such demonstrations; i.e., those in which useful results were competently reported to decision-makers who could then act on this information. Can you give some examples of "useful" experiments?
A. It is, for a variety of reasons, easier to measure the results of an accident countermeasure when the change is rather direct; i.e., when some component of the system has been changed which is very directly associated with the accident situation. Consequently, it is easier to find examples in the field of highway engineering, with the direct effect on the traffic stream, or in vehicle safety engineering and the direct effect in injury reduction. In answer to this question we take several examples from the highway field.

The San Diego Pedestrian Study. The San Diego, California, Traffic Engineering Department (2) recently completed a program under "402"* sponsorship which illustrates one possible outcome--a change (countermeasure) which was expected to be beneficial, but in fact was not. A study of pedestrian accident frequency in and out of painted crosswalks indicated that such accidents occurred more often than expected in crosswalks. Further observation led to the suspicion that San Diego pedestrians felt safe in crosswalks because they knew they had a legal right-of-way but that automobile drivers could not see the crosswalks, with their shallow vision angle, nearly as well as the pedestrian thought they could (or as well as he could see it).

In this case the Standard had suggested, although not explicitly stated, that more painted crosswalks were good for people; and the analysis raised some doubt. The result of the study was a modification of the warrant for painting crosswalks in San Diego; new ones evidently will not be installed indiscriminantly.

Head-On Collisions on Expressways. Several busy expressways have had severe problems of head-on fatal collisions. Before median barriers were installed on California freeways, 20% of the fatal accidents were head-ons. After installation of a barrier where there was a narrow median (over 400 miles of barrier) this figure dropped to 4%(3). Similar experience was had in New Jersey(4); fatal head-on collisions practically disappeared. The positive results were well reported and were no doubt responsible for the other states taking similar measures.**

The New Jersey Sign Changing Experiment. In 1970 in New Jersey a signing change on Route I-287 at its intersection with US-22 was intended to reduce the frequency of turning errors committed by drivers at that point. This in turn was expected to reduce the probability of an accident at the same location. A careful count revealed that the error rate dropped from 5 per hundred vehicles to $4\frac{1}{2}$ per hundred--statistically significant with the amount of data taken, but of somewhat doubtful practical value in this particular case. Still it informed the engineers of a better choice of signing in their continuing programs.

2.1 THE GENERAL CASE

But as we depart from the cases which have been influenced directly, it is often not as clear that a change has been effective. The relationship between a change (say a new simulator for the driver education classes in Idaho) and the reduction of accidents in that jurisdiction is often difficult to define. In fact, the whole purpose of the study being reported here is to get at that kind of problem.

Q. Who uses the results of such evaluation programs?

A. The results of an evaluation should be useful at several levels. Consider the diagram shown in Figure 2.

Let's assume that project "A" is being conducted in a state highway department, perhaps to develop a staff and a technique to identify the most hazardous locations in each highway district. The results of this program are certainly of interest to the project director--he should want to know on a day to day basis whether his work

*"402" projects are programs conducted within the state using matching federal and local funds. These are colloquially called "402" programs since the authority to conduct them derives from Section 402 of the Highway Safety Act of 1966.

**Incidentally, the need for a "controlled" experiment disappears when such massive results are available. But this is a relatively unusual case.

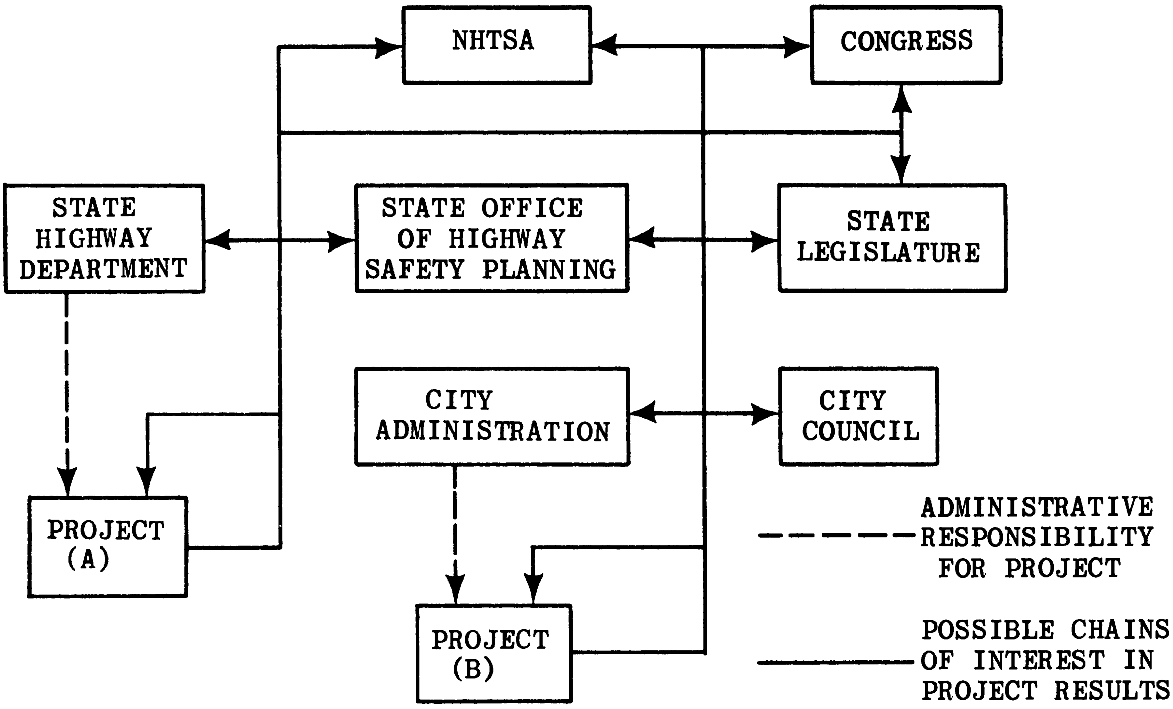


FIGURE 2. THE INTEREST IN THE RESULTS OF PROJECT EVALUATION

is successful. But it may also be of interest to the highway department, which will be involved in the decision of whether or not to fund this activity in the ensuing years. And the state office of highway safety planning should have an interest, because this may be a program which could be "sold" to county level organizations on the basis of success at the state level. Further, the legislature may be interested because it approves the highway department budget, and the National Highway Traffic Safety Administration or Federal Highway Administration should have an interest so that the program could be replicated in other states.

Similarly, if a program is conducted at a city level (say to improve its ambulance service), the results may be of value to everyone from the project director up. In fact, the only time it is not appropriate to find out whether a program was effective would seem to be where there was only one possible outcome--success. This is indeed the case in some programs, but more often than not someone will want to know the measure of success.

Q. This scientific approach to evaluation doesn't seem very useful. Isn't it better just to look at the ultimate outcome; i.e., to do a "practical" evaluation?

A. There seem to be a lot of terms used to talk about the evaluation process--practical vs. scientific, ultimate measures vs. intermediate or proxy measures, efficiency vs. effectiveness, etc. These all deserve some explanation so that we can understand each other.

There are two rather diverse approaches to getting a measure of the value of some countermeasure. They are perhaps well represented by the viewpoints of Walt Whitman on the one hand and Professor John L. Synge on the other.

When I heard the learn'd astronomer,
When the proofs, the figures, were ranged in columns
before me,
When I was shown the charts and diagrams, to add,
divide, and measure them,
When I sitting heard the astronomer where he lectured
with much applause in the lecture-room,
How soon unaccountable I became tired and sick,
Till rising and gliding out I wander'd off by myself,
In the mystical moist night-air, and from time to time,
Look'd up in perfect silence at the stars.

Walt Whitman, "When I Heard the Learn'd Astronomer"

Were Walt Whitman responsible for making budget allocations in the traffic safety field it is not clear what his method might be. It is cheap to look at stars, or perhaps to listen to music, or to feel the warm wind--but it is expensive to pay more policemen, or to reconstruct old roads, or to have a system of vehicle inspection. And it seems unreasonable that such decisions should be made primarily on the basis of feeling.

> It is all very well to say that the world of reality should be kept separate and distinct from the world of mathematics. In the trivial operations of daily life, you may be able to keep clear of the concepts of mathematics, but once you begin to touch science, the dangerous contact is established.
>
> John L. Synge, "Science, Sense and Nonsense"

On the other hand the pure statistician is often a very rigid person. He may insist on a perfect experiment with appropriate controls chosen with attention to randomization of the factors, and by his rigidity insures that the experiment cannot really take place.

So here is the quandary--shall we just go out and "do our thing" without any measure other than the public's eye; or shall we conduct a highly controlled experiment from which we can infer conclusions which are not very useful to anyone? The answer would seem to be neither--it would be much more helpful if we could operate in such a way as to get a reasonably acceptable measure of a program's effectiveness in terms understandable to the people who make the decisions as to what to do next.

2.2 PRACTICAL VS. SCIENTIFIC APPROACH

When a program is undertaken the sponsor may have two questions in mind: (1) did you do what I told you to do?; and (2) did doing it do anybody (i.e., the public) any good? The first of these is often thought of as a practical evaluation and indeed it is the basis for much of the NHTSA's current evaluation manual--getting some measure of whether the states actually undertook the programs required by the Standards themselves. The second, however, is often thought of as the scientific evaluation, and it should be obvious that both words, practical and scientific, are misnomers. Both kinds of evaluations can be either quantitative or precise; both can be extremely useful. The second question is particularly useful with respect to innovative programs where the outcome is not obvious. Further, the second leads to considerations of cost effectiveness or cost benefit, which could be helpful in deciding whether to continue the particular kinds of work.

Q. What do you mean by cost-effectiveness and cost-benefit?

A. These terms are sometimes used rather loosely, and we will attempt to define them for our purposes here. Both have to do with the second type of evaluation discussed in the previous question, i.e., did this program do any good? However, cost-effectiveness often cannot be put in terms of dollars, but must instead be stated in terms of some accomplishment level per dollar expended. For example, it seems appropriate to consider the cost of a new driver simulator, and to compare the number of students taught (per dollar expended) before and after the simulator was purchased--perhaps just assuming that the educational value of the simulator was at least as great as the system it replaced. Such measures are sometimes called intermediate or proxy measures as they serve in lieu of some ultimate accident measure.

From a cost-benefit computation point of view, we are better off if we can take some measure of the ultimate effect--lives saved per dollar expended, injuries prevented, accidents avoided, etc. Similarly, from the point of view of getting some governmental body to adopt a program, we would be better off to have our results in a form as close as possible to the ultimate goal. But we are often just not in a position to do this. For example, the relationship between vehicle inspection and the number of fatal or other accidents is very tenuous. It seems likely to us that there should be such a relationship, but we just do not know how yet to get at it quantitatively. So we are content with discovering whether a motor vehicle inspection system will decrease the number bad brakes, broken windshields, loose steering systems, etc., on the premise that these defects may contribute to accidents.

In each highway safety field there is a sequence of measures ranging from the immediate effects of a countermeasure to the ultimate. In driver education an immediate effect is the communication of knowledge to the student, and this can be measured by an examination. We make the assumption that unless the driver knows, for example, where the brake pedal is, he probably will not be able to stop the car. A measure closer to the real problem would be to take the driver out on the road and ask him to stop the car and to score him on his ability to do this. And, from a safety point of view, the ultimate measure would be whether or not the driver gets into or causes any accidents once he enters the traffic stream.

Our task in the present program has been to design the data taking and analytical methods of each demonstration program to obtain information which will be of greatest value to the people who must make decisions regarding implementation of future programs. For those countermeasures which are close to the accident process, what we have called direct component changes, the measures are often in terms of accident or injury reduction. But for other programs, changes in the driver education curriculum, for example, the measures may be in terms of road performance or even knowledge and attitude when that seems the most useful.

It would be so much more advantageous if one could compute the dollar value of the different methods of driver education (in terms of accidents saved over the lifetime of an individual's driving) and in some instances it may be possible to do this. If one could put a dollar value on the accidents prevented, and compare that with the cost of the training programs one could arrive at a pure cost-benefit ratio. As we have noted, it is easier to compute the benefits if the program is close to the accident process.

Q. But intermediate or proxy measures don't necessarily indicate that a reduction (in accidents, injuries, or fatalities) will follow. For example, you said that painting more crosswalks in San Diego was a bad thing; if I had just counted painted crosswalks I would have assumed success when in fact things were worse. Isn't this another quandary?

A. It certainly is. There is an old tale about the horse who lost a shoe in battle--"for want of a nail, the shoe was lost, for want of a shoe the horse was lost, for want of a horse the rider was lost, for want of the rider the battle was lost, and for want of the battle the kingdom was lost." And it seems that there is a fairly logical relationship. We could certainly institute a quality control program in the nail factory to insure that we had better nails for our horseshoes, but it seems unlikely that we would insist on measuring the effectiveness of this program by counting lost kingdoms.

Nevertheless it would be equally invalid to measure some intermediate variable which did not have logical relationship to the ultimate outcome, say the birth rate of horses. But it is obvious from the San Diego pedestrian case that we should be more than careful in establishing the relationships.

A possible problem in measuring the effect of some change is obtaining a sufficient quantity of data. Lost kingdoms might in fact be appropriate if there were enough of them. But there are likely to be plenty of nails, and perhaps also horses and riders, to be statistically useful. A change in rider losses then, might be at the same time an indication of the efficacy of the nail quality control problem, and some measure of the probability of losing the battle and thence the kingdom.

Experience indicates that when comparing some observed quantity before and after a change has been made, it is important to decide whether or not the observed difference is too large to have occurred by chance. A somewhat conservative test often used is called the chi-square, which permits the observer to decide whether a change in the number of accidents or other phenomena could have happened by chance, or whether there must have been some causative factor. A simplified version of this test has been plotted in Figure 3, adapted from Van Vechten's study (5) showing the relationship between the number of events (accidents) before and after some change has been introduced. To use the table one enters with the numbers for the before and after periods, locates their intersection, and concludes that (1) there is enough data to indicate that there has been an improvement; or (2) there is enough data to indicate that conditions are worse; or (3) there is not enough information to decide one way or the other.

More data may later permit decision (1) or (2) to be reached; whether one should take more data (over a longer period of time) will depend on how badly he wants to know whether the change was effective.

Note that fairly large numbers are required for the experimenter to be at least 95% sure that the difference observed is not due to chance but rather to causative factors. In the traffic field one will often not be able to count accidents for a long enough period to get the required data. If a single intersection is changed, and the number of accidents goes from 30 in the before to 20 in the after years, no conclusion can be drawn. But if one could have measured driver errors which are believed to contribute to accidents (say wrong left turns or sudden stops), a significant reduction in these might be observed in a few days. The ability to decide within a few days whether some change has been produced should complement the ultimate measure of accident reduction.

Q. How does one decide whether a demonstration or an experiment has been successful?

A. It has been said that statistics is the art of drawing conclusions in the face of uncertainty; and it is for this reason that statistics

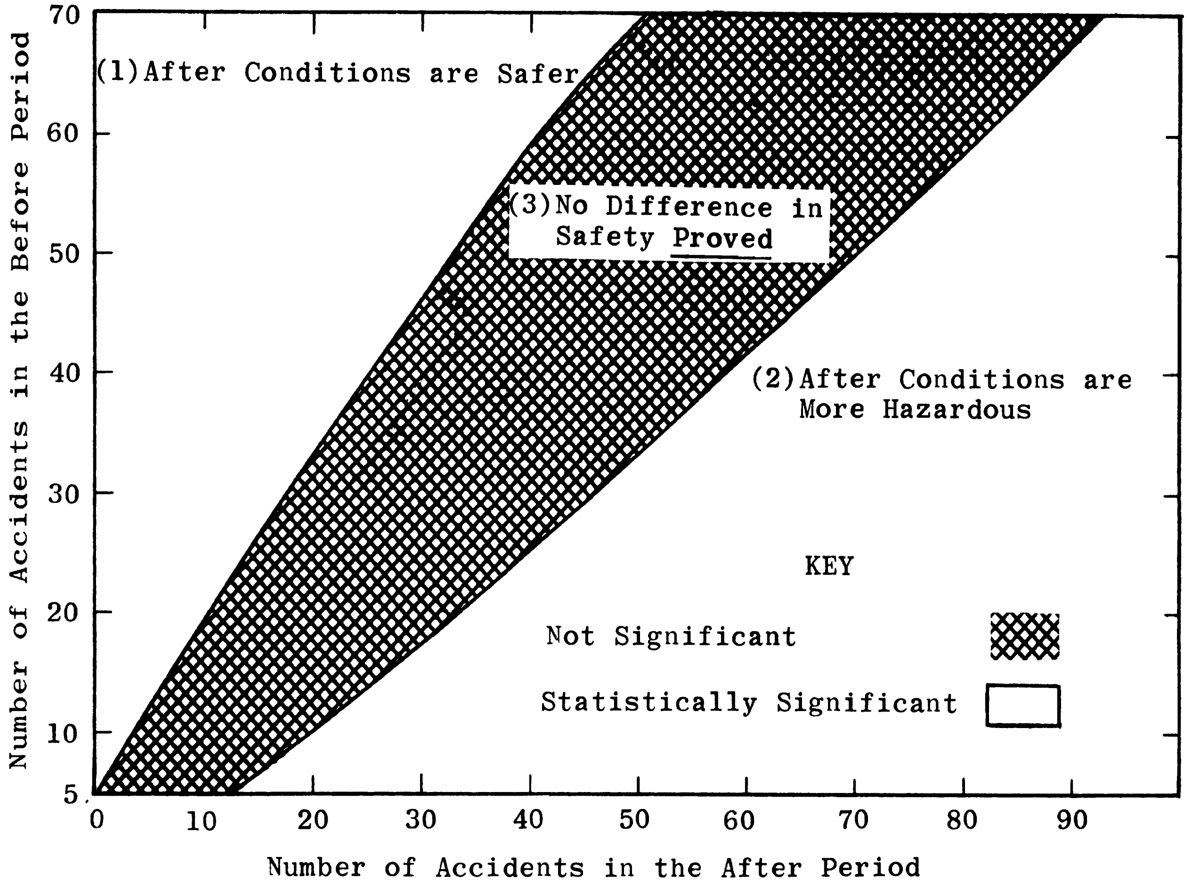


FIGURE 3. CURVES FOR DETERMINING IF A DIFFERENCE IN THE NUMBER OF ACCIDENTS IS STATISTICALLY SIGNIFICANT BY CHI SQUARE AT 95% LEVEL OF CONFIDENCE

is called upon in the evaluation process. In statistical terms we ordinarily design an experiment, take a set of data, analyze it, and draw some conclusion about the outcome.

B.J. Campbell(6) has articulated nicely the possible outcomes of an experiment, and his comments are paraphased here. We can look carefully at the data, and then conclude that the program was a success, that it failed, or that we do not really have enough information to claim either success or failure. There is, of course, a possibility of error in the judgment--we could conclude that there was a success when in fact things were not any better (this is called a type I error in statistical terms); or we could conclude that the program was not successful when in fact it was (this is called type II error). Campbell points out that scientists abhor the type I error; they would hate to announce a new theory when it was in fact incorrect. Practitioners,* on the other hand, abhor the type II error; they would not like to decide that a program was unsuccessful when in fact it was useful. And since the probabilities of each error type are interdependent there is a problem.

Statistics, however, furnishes a language through which the scientist and the practitioner can talk with each other in spite of their differences. And hopefully the precise language of statistics will allow everyone to understand the results of an experimental or demonstration program.

Q. How does the practitioner learn the language?

A. It would seem that the practitioner should somehow obtain a prac-

*For example, medical researchers hesitate to declare a new drug ineffective if there is any shred of hope.

tical understanding of the field of statistics. This may be a little frightening at first thought, but we suggest that it is the better of two choices. For example, one could ask a statistician to learn enough about the field under investigation (driver education, police traffic services, highway engineering) to conduct the design and evaluation of an experiment; or one could ask a practitioner (a highway engineer, policeman, teacher) to learn enough about statistics to do the same thing. The latter approach is certainly being used in the highway engineering fraternity, with more and more traffic engieers receiving a modest amount of training in experimental design.

One may still occasionally wish to discuss a problem with an expert in the field of statistics, just as occasionally it is useful to call in a consultant on some other aspect of a program. But a little training in the experimental method will provide a capability within an operating agency, (as in the traffic safety field), for both solving problems and for understanding and communicating with others about the problems.

The present report is, of course, not a textbook on the subject of experimental design; but our experience has been that a single three hour per week 16-week course in statistics can get the average practitioner to the desired level of competence.

Q. What are some of the problems which require an understanding of statistics?

A. There are a number of statistical dangers confronting those conducting and evaluating an experimental program. We will discuss here only two: (1) the law of averages, and (2) the need for a control against which to compare the experimental results.

Q. What does the law of averages have to do with evaluation?

A. The law of averages is given a much more sophisticated name by statisticians--the regression to the mean. It is such a simple and important concept, that it seems to be worth a more complete discussion here.

Figure 4 represents a sample of drivers characterized along the horizontal axis by how well they behaved during the past year. Those at the left of the plot were the best drivers of all; they not only had no accidents and no violations, but they were completely error-free. They never made wrong left turns, exceeded the speed limit, failed to stop at a stop sign, etc., for an entire year. Of course, there were not many of these drivers as indicated by the small shaded area at the left.

On the right side of the figure is another shaded area, which includes all of those drivers who were responsible for fatal accidents during the past year. On this particular scale they are considered the worst drivers. In between are a small group of drivers with three violations and one accident (they are somewhat worse than average), and a group with some (say 100) errors, but no violations or accidents. The peak of the curve, which is symmetrical, indicates the average driver, and most drivers in this example are seen to be near average.

In this case the expression "regression to the mean" means simply that in the next year it is unlikely that all of the worst drivers (those responsible for fatal accidents and still living) will stay in that group, for some will move or regress to the mean or average. They may still be bad drivers, but probably very few of them will be involved in fatal accidents again. So if this group is measured next year and their position computed on the goodness-badness scale they might, on the average, be in the same position as those people who had three violations and one accident the year before. They could thus have regressed (moved) toward the mean or average (the center of this distribution).

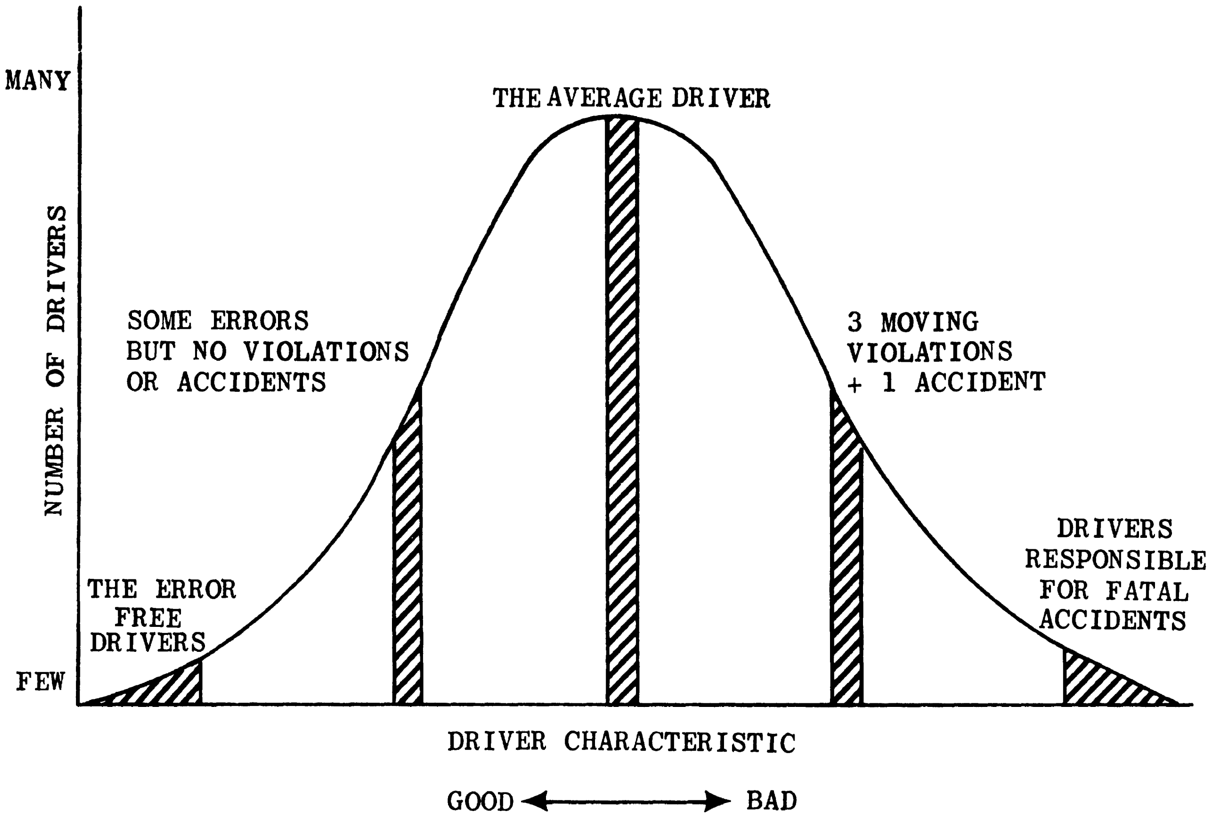


FIGURE 4. THE CHARACTERISTICS OF DRIVERS IN A HYPOTHETICAL SAMPLE.

Drivers who have been responsible for a fatal accident and lived have had a rather traumatic experience; is it therefore reasonable to conclude that this experience caused them to become better drivers the following year? Certainly they did not stay where they were at the worst end of the scale. Or is it just that the worst drivers will get better by the law of averages? Some of them were there because they truly were bad drivers, and they may continue to be bad, but some may have been good drivers who had bad luck and they may be on the other end of the scale next year. If there is any chance involved in their position this group will, on the average, get better.

Now let's develop the same argument about the small group which had three moving violations and one accident. On the average are they likely to be better or worse next year? Or are they likely to stay the same? Or how about the group who made a few errors, but had no violations or accidents--is it possible that they will have some accidents next year?

Again, if chance has been involved in some way in getting them to their present state, it is likely that the good will get worse, and the bad will get better. The implication of this phenomenon to experimental design is that if we simply take the group of drivers who had three violations and one accident, force them to take a remedial course in driving, and then find their performance during the following year better we will not really be able to decide whether the change was the result of the course or the result of the law of averages being applied.

Q. What can be done about this?

A. That is where the use of a control group comes in. Of course if we could be sure at the outset that the experimental group was average, as they might be if they were selected at random from a group of employees without regard to their driving record, then we could conclude that a change in their performance was in fact the result of the course. But in the more usual case it will be appropriate to have a second group similar to the first, chosen using the same criteria, but to avoid giving the second group the course. At the conclusion of the experiment the performance of the two groups could be compared,and any difference appropriately assigned to the course itself.

There is often a great hesitation to create a control group, usually because the program manager is personaly convinced as to the value of the treatment (e.g., the course) and he would not think of letting anyone not take it. E.B. Wilson (7) tells the story of the African pygmy tribe which is convinced that the beating of tom-toms brings back the sun after an eclipse. They have never run the control experiment and they probably never will. And while you laugh at the pygmies, think also of the governor who introduced the speed crackdown in a high accident year, and claimed full credit for the drop in accidents the next year. This is not to suggest that there may not have been an effect, but merely that the combination of the law of averages and the lack of a control for the experiment puts him in about the same position as the pygmies.

Regression to the mean and the need for controls apply to larger scale problems, too. If one selects the 10 worst cities, based on the fatality rate, for change programs, and then treats all of them, it can be expected that they will, just like the very bad drivers, get better. It will be difficult to assign the reason for improvement to the change program unless some comparative controls are available.

Q. Sixteen standards have been promulgated to the states...why don't the states and the subsidiary jurisdictions just go ahead and do what they were told?

A. There still seem to be a lot of people as well as government agencies who insist on proof--they want to be shown the value of the recommended programs. The answer would seem to be to demonstrate the effectiveness of the various countermeasures, measure the results, and present the evidence to the disbelievers.

Q. But it is not possible to measure everything, is it?

A. No. Someone will have to choose which areas are most important and do these quite carefully.

Q. How should these program areas be chosen?

A. There doesn't seem to be any very unique method of making the choice, but we suggest a method here which seems to have some logic behind it. It is based on the idea that certain groups of change programs suggest a common set of measurements, i.e., they operate in ways which are similar enough that they could logically be combined into demonstrations of effectiveness.

Table I shows the 16 current Highway Safety Program Standards in their usual order at the left of the diagram. The second column indicates that there are many invididual countermeasures including such elements as "send one police officer to a nine-month training course," or "begin a driver education program in the rural area schools" or "install a 360/50 computer at state police headquarters to process accident data." Each of these elementary countermeasures is a possible change program in any area.

We have grouped these elements into six program areas shown in column 3. They are not necessarily in direct correspondence to the Standards of column 1, but for the most part the relationship is obvious. The arrows show the links for an emergency medical program.

TABLE I. STANDARD, COUNTERMEASURE, AND PROGRAM AREAS

HIGHWAY SAFETY PROGRAM STANDARDS	ELEMENTAL COUNTERMEASURES	DEMONSTRATION PROGRAM AREAS
1. Periodic motor vehicle inspection	1. Buy a radar	1. Vehicle regulation
2. Motor vehicle registration	2. Train a medical attendant	2. Driver regulation
3. Motorcycle safety	3. Enact a law requiring helmets	3. Environmental regulation
4. Driver education	4. Inventory the traffic signs	4. Driver preparation
5. Driver licensing	5. Create a medical board for license review	5. Information flow
6. Codes and laws	6.	6. System restoration
7. Traffic courts	. etc.	
8. Alcohol in relation to highway safety	. etc.	
9. Identification and surveillance of accident locations	. etc.	
10. Traffic records	.	
11. Emergency medical service	.	
12. Highway design, construction, and maintenance	.	
13. Traffic control devices	.	
14. Pedestrian safety		
15. Police traffic services		
16. Debris hazard control and cleanup		
17. Accident investigation		
18. Youth transportation	1000's	

We have viewed the highway traffic system as consisting of two loops, both proceeding from the occurrence of some event. The event most easily considered is an accident, although the model can be extended to permit inclusion of violations, or even administrative events such as license renewals. Consider first the case of an accident.

When an accident occurs the two loops are set in motion. The short term loop involves the immediate detection of the accident, and the necessary responses--ambulances, tow trucks, debris removal equipment, and the loop closes when order is restored to the site. The long term loop begins with the reporting of the accident to some information storage medium such as a newspaper, a pin map, a computer, or by word of mouth to the chief of police. But whenever the information is stored we assume that sooner or later there is going to be some analysis of the information. Again it may be formal or informal--the traffic officer may simply note a black spot on the pin map, or a large computer may spit out the fact that the accident rate at this intersection is double that of the year before.

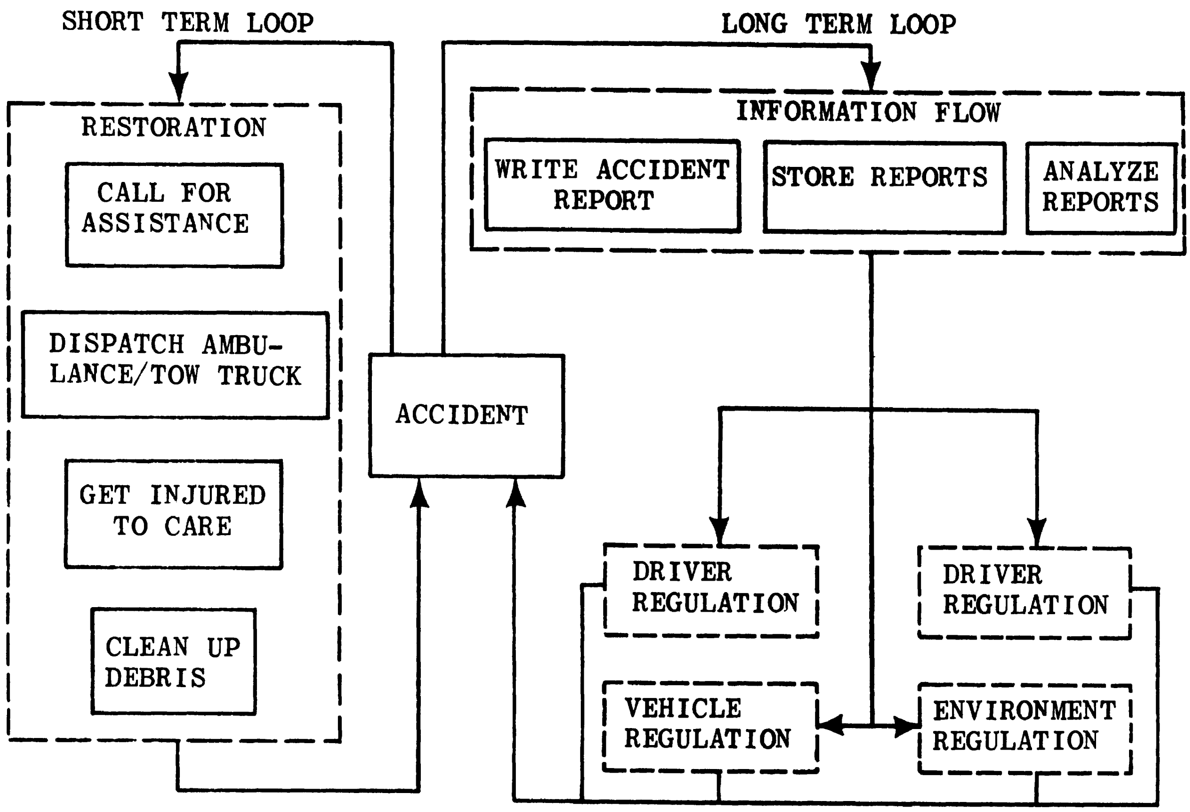
In either case the results of the analysis may induce some agency to take action, perhaps by installing a traffic signal, repainting an old sign, etc. The effect of the action, then, is to change the characteristics of the highway (environment) in such a way that its performance is modified. Traffic operations through that intersection are henceforth different than they were; hopefully different in a direction which will reduce the probability of an accident at that point in the environment. Thus the long term loop closes back to the event.

By extension events which set the loops in motion could include other sorts of information producers--violations in which the information becomes a part of the driver record, or vehicle registration (the information becoming a part of the vehicle record). The analysis could be done by a variety of agencies (highway, education, manufacturer) and action in general will be taken with respect to either the environment, the vehicle, or the people (drivers, pedestrians) in the traffic system.

In Figure 5 there are dashed lines drawn around the parts of the highway traffic system which represent the six demonstration program areas shown in Table I. They have the particular virtue, from an experimental point of view, that they permit grouping of countermeasure (change) programs at the same location in the traffic system, implying that the kinds of measures to be made are common within each grouping.

Q. There are six program areas, but they obviously could incorporate many, many individual change programs--which should be chosen and how should the choice be made?

A. By way of example consider the area of emergency medical care. Possible actions we could take in this area are that we could: (1) provide further training to the attendants; (2) obtain new medical and vehicle equipment; and/or (3) enhance the communication system and the organization (management) of the emergency medical force in the jurisdiction. In fact, essentially all of the individual change programs suggested in the federal Standard relative to this area could be classified into one of these three categories. But it would also be quite possible to do more than one of these things at the same time and in the same place, or in fact to do all three. Further, any one of these changes could be made at any of several levels of intensity or complexity. For example, one could train the ambulance personnel to the level of a beginning Red Cross course (perhaps the lowest level), to the NHTSA standard course (perhaps an intermediate level), and to the level of a military corpsman or even an intern (the highest level).



Restore situation to normal Reduce the probability of another accident

NOTE: The dashed-line boxes indicate the six experimental programs defined in this report.

FIGURE 5. THE SHORT AND LONG TERM LOOPS ASSOCIATED WITH A TRAFFIC ACCIDENT

All of the possible combinations of these actions are shown in Figure 6. There are a total of 21 blocks which constitute reasonable experimental cells; that is, they define reasonable activities for change programs which might be measured carefully to find out how much good they really did.

Within the spirit of the SCOPE program there is clearly not enough funding for 21 experiments, even if they were desirable. Each program area will have its own peculiarities, however, which may suggest which of these cells are most important. We suggest here some of the subjective reasoning which goes into choosing a program--i.e., which experiments should be done first; which have the highest priority.

From the joint point of view of the experimental designer and the administrator at the federal level, we should seek those programs which:

(1) Can be funded; i.e., will Congress put up the money for an experiment or demonstration in this area?
(2) Exhibit a clear need for more information. There is not much point in demonstrating in some area in which everyone (or almost everyone) agrees that it is a good idea.
(3) Have a reasonable chance for a successful demonstration. There should be an identifiable goal with some anticipation that it can be reached and measured.
(4) Can interest someone or some jurisdiction to conduct the demonstration or at least to cooperate.

With these considerations in mind we might conclude that there is little to be gained by an experiment in which training alone is varied. Further we might decide that it would be most appropriate if we could enhance all three areas (training, management, and equipment) and try to measure in a single demonstration the effects of each independently and in combination.

The three lower blocks in Figure 6, then, would receive top priority in planning; the 18 upper blocks would be considered "nice to have," but not mandatory. It seems likely, however, that with the many states implementing "402" programs that there will indeed be natural experiments occurring and that we should plan to monitor the appropriate "402" programs carefully to take advantage of these.

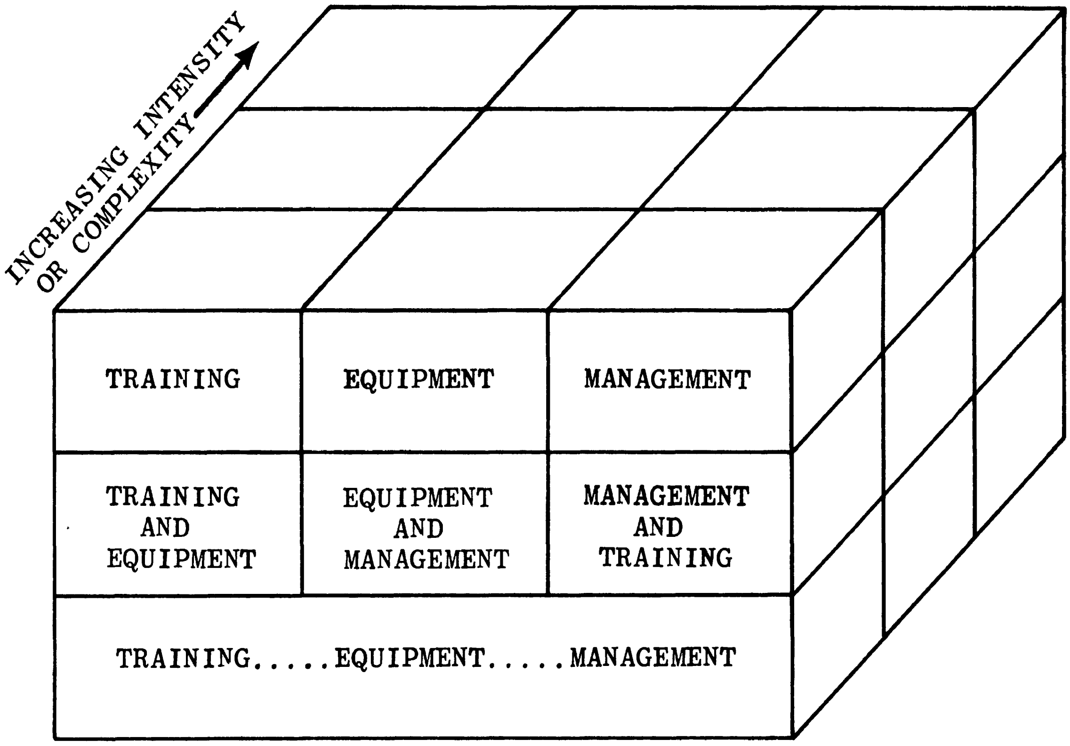


FIGURE 6. POSSIBLE EXPERIMENTAL CELLS FOR AN EMERGENCY MEDICAL COUNTERMEASURE PROGRAM.

Q. How does one choose sites for demonstration programs?

A. Generally, sites for experimental or demonstration programs should have the following properties:

(1) There should be a problem to be solved; e.g., there is little value in trying to improve an adequate ambulance system.
(2) The jurisdiction must be large enough for valid experimental results. That is, we must be able to collect enough data to be able to say with some degree of assurance that things were indeed better as a result of the demonstration.
(3) The jurisdiction must be hospitable toward the demonstration effort, particularly toward the experimental and measurement aspects of it.

Q. How much will all of this cost?

A. We have suggested, for each program area, a number of experimental cells which are considered to be of first priority in the spirit of the SCOPE program; i.e., obtaining demonstrable evidence as to the value of countermeasure programs recommended by the Standards themselves. We note, however, that establishing more or fewer programs of this kind is a matter of judgment and is obviously dependent upon the available budget. With all of these uncertainties we have defined experimental programs which are discussed in later sections of this report. We believe these will provide the desired evidence.

It is useful to ask the question in the form: What value does this demonstration program have? The purpose of a demonstration program is to convince someone (a mayor, a state legislator, etc.) that he should commit funds to make some change in a highway or traffic system. And he should be convinced that the change will be of some value to the public.

Detailed cost information for each recommended program is given in the appropriate section of Volume III of this report. A brief summary is given in Table II, and where a range of costs is given, indicates that several alternative program implementation levels are discussed. It seems likely that many of these programs will be funded in part by state and local monies. Costs shown in the table are given on the basis that the programs would be fully funded by federal sources, and should be modified downward depending on the degree of local support.

TABLE II. ESTIMATED EXPENDITURES FOR COUNTERMEASURE DEMONSTRATION PROGRAMS

Program Category	Period (Years)	Operation[a]	Instrumentation and Evaluation[a]	Total[a]
Road User Regulation	3	1.35 to 2.25	1.35 to 1.95	2.7 to 4.2
Information Flow	3	0.51 to 2.85	0.1	0.61 to 2.95
Road User Preparation	3	0.75 to 1.2	0.42 to 0.54	1.17 to 1.74
Vehicle Regulation	3-5	5.9 to 13.1	0.1 to 0.9	6.0 to 14.0
System Restoration	3.5	4.0	0.33	4.33
Highway Regulation[b]	---	-----------	-----------	-----------

(a) Indicated program costs are in millions of dollars.

(b) In the Highway Regulation area no specific operational program is proposed, although an example of the design for an experiment is given. No costs have been presented for this.

3. SUMMARIES OF SIX DEFINED DEMONSTRATION PROGRAMS

This section of Volume I presents summaries of the six experimental programs which have been defined as a result of this study effort. These include, in order: (1) road user regulation,(2) information flow, (3) road user preparation, (4) vehicle regulation, (5) system restoration, and (6) highway regulation.

Each of these experimental plans is covered in greater detail in Volume III of this report. Readers with a particular interest are encouraged to refer to that volume.

3.1 ROAD USER REGULATION DEMONSTRATION PROGRAM

Presently there seems to be a need for evidence concerning the value of countermeasures in the several fields associated with law enforcement. This is stated as an hypothesis, and there are certainly those who would disagree.

In normal operations in the law enforcement field anecdotal evidence is widely accepted and indeed many changes are introduced on the basis of such evidence. A city may buy and operate traffic radars, change the court system to referees instead of judges or vice versa, or pass new traffic laws (such as "if a vehicle is traveling so slowly as to hold up more than five other vehicles on a public highway he must pull to the right and let the following vehicles pass"*). Not only will such changes be made, but they will be recommended to others by those who adopt them, and they will propogate. There are few statistically meaningful studies of the efficacy of law enforcement activities; of those done, one of the most well known is California's Operation 101 which has developed experimental evidence of a relationship between enforcement and accident reduction. While this study found the direction of the relationship, it was not adequate to define it quantitatively. A more recent California study, Operation 500, is attempting to do that.

Yet, many people remain unconvinced of the value of specific people regulation countermeasures, particularly the city councils and state legislatures who vote the dollars for supporting the police and the courts. While no one seems to doubt the need for some degree of law enforcement activity, there is often much hesitation about allocating money for changes. For example, the city of Flint, Michigan, coincident with the operation of a strong federally supported traffic enforcement program, enjoyed a reduction in the number of fatalities from 31 in one year to 13 in the next. The operating agencies are tempted to give the credit for the change entirely to the law enforcement program, and at a currently accepted value of $140,000 for a life lost in traffic, the change should be worth $2,500,000 annually. Expenditure, on the other hand, for the new police service was less than $400,000 in one year. But police administrators are having a difficult time convincing the city of the program's value.

So here is the problem. If, as we expect, there are a variety of law enforcement practices which are worth more than they cost we would be well advised to collect proof of this in such a way that we can convince the city councils and state legislatures. This section of the report addresses the problem of collecting such evidence.

We have placed most of the accident countermeasures having to do with the dynamic regulation of people--primarily drivers--into one group for discussion purposes. This is both convenient and useful from

---

*A Michigan statute.

the experimental design and measurement points of view. Included in this grouping are change programs suggested under the federal Standards regarding Police Traffic Services, Codes and Laws, Courts, and Alcohol, and part of the Pedestrian Standard. Many individual countermeasures are possible, of course: add one new judge to the court systim, change the procedure for sanctioning traffic violations to an administrative fine, buy a Vascar for the local police department, etc. Since someone can be found who will strongly support any one of these, they may all be candidates for experimental study.

We have considered an experimental approach to get data regarding the value of such countermeasures in a useful form. We would like to at least aim in the direction of ranking the payoff of the several possible countermeasures in this field, and ultimately would like to get at the problem of comparing their value with that of countermeasures in other areas. We have noted earlier that this is a difficult task, but one worth striving for.

For discussion purposes we have grouped these individual countermeasures (things to be done), into eight areas which also are useful for experimental implementation. The eight areas shown in Table III might be implemented singly or in any number of combinations. Assuming that one could implement each possible combination at three intensity levels, there would be 747 cells. Many cells may actually come to exist in programs throughout the country, but we do not believe that such a large number of experiments would be a fruitful goal in a set of demonstration programs. We do recommend, however, that some of the combinations be tried in the context of current demonstration programs and that, at a minimum each countermeasure area be singly implemented in at least one jurisdiction. Four of the most often used change actions--media, training, manpower, and social handling--are examined at length, and a specific program involving all four categories is discussed in the context of a particular city, Chattanooga, Tennessee.

We recommend further that two cities or counties be selected for concentrated long-term experimentations. In these locations new techniques and devices, or operational changes in law enforcement can be introduced from time to time and their effects can be observed through a rather complete set of instrumentation. These two jurisdictions would constitute a national laboratory, guided by a senior advisory council of persons with national prominence.

The change process operating in the national laboratory areas will be similar to that followed in most communities. The distinct phases of problem identification, solution formulation, implementation, and evaluation will be followed, but unlike the typical community these activities will be guided by advisors who are among the most capable individuals available, and each step will be carefully measured to demonstrate its influence on problems. Two types of change will be studied: major system alterations, such as a substantial increase in manpower, and refinements of specific techniques such as assessing the relative effectiveness of radar, Vascar, and Orbis in detecting speed violations under particular conditions. The changes will be implemented sequentially after careful consideration of which activity will yield the greatest incremental benefit. Results of these changes will be measured and analyzed; the most cost-effective will then be recommended to other communities.

We suggest that these two jurisdictions be Bloomington, Indiana and the surrounding Monroe county, and Flint, Michigan, perhaps again including the surrounding county. Bloomington has had a continuing set of instrumentation designed to unobtrusively sample traffic activity (speeding, etc.) in Monroe County. This instrumentation has

TABLE III. ROAD USER REGULATION
COUNTERMEASURE GROUPS

1. Management improvements are overlaid on a ongoing Police Traffic Services (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 a specific tool (e.g., 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 and laws where streamlining, standardizing, and training occur mainly dealing 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 (alcoholics).

been developed to a relatively high level by Indiana University personnel, and provides a most useful test base. In addition there is a small number of separate police forces in this county, and they have shown a continuing interest in participating in experimental studies.

Flint has had a strong program of traffic enforcement for several years, and has developed a large traffic unit within the police department with specific selective enforcement duties. While their instrumentation at present is not as sophisticated as that of Bloomington, they have been collecting a variety of data relevant to the evaluation process. In addition we believe that instrumentation similar to that in Indiana could be installed at modest cost, giving an opportunity for continuing unobtrusive monitoring of traffic activities as a function of changes in the enforcement system.

Measures of effectiveness to be taken in connection with these programs would include three levels: a comprehensive measure of activity as determined by a continuing recording of personnel performance; a set of intermediate measures including changes in speeding, changes in the attitudes of the drivers; and ultimate measures (accident, injury, and fatality counts). There will be a stress on the intermediate variables as the current measures of success or failure of parts of the program, but an important part of the national laboratory effort will be to link these intermediate measures with the ultimate goals.

No detailed cost estimates have been prepared for these programs. In the four-activity experiment outlined in Volume III evaluation expenses will be on the order of \$10,000 to \$12,500 for examining three activities in detail and from \$30,000 to \$60,000 for recording and evaluating changes in violations over three years. The total costs of the national laboratories, including increased police manpower, will be $500,000 to $750,000 per year. The national laboratory should be planned for five years with possible indifinite continuation if it proves successful.

This entire program is laid out on the assumption that useful information will result. We expect that there will be more solid and convincing evidence regarding the various facets of the law enforcement process--ultimately leading to better decisions by the money providers, and thence to reduced carnage on the highways. We are not particularly hopeful about developing full cost-benefit information within this field, but we do believe that the approach outlined will provide decision-makers with sound and understandable evidence about the value of the many possible countermeasures discussed.

3.2 THE INFORMATION FLOW DEMONSTRATION PROGRAM

While the majority of activity concerned with the flow of information in the traffic safety area has been identified under the Traffic Records Standard, there have been a number of additional federal-state matching fund programs which involve the handling of data. In the first three years of "402" funding approximately $37,000,000 of federal money was allocated to Standard area 10--Traffic Records. But in a study of all "402" programs, more than $50,000,000 was identified with similar work (e.g., computer record keeping), including programs funded under motor vehicle regulation, identification and surveillance of accident locations, and police traffic services. The majority of these funds went to create a capability for the automatic processing of traffic information: accidents, driver records, vehicle records, highway records. But to date this capability seems to be much underused.

It is clear that the intent of the Traffic Records Standard and others indicated here was to promote the ultimate use of processed information by both state agencies and local jurisdictions within a state. A rather secondary purpose must have been to develop data which would be of statistical value to the federal government in evaluating its efforts or in identifying new problem areas. We believe that a demonstration program in this field, oriented particularly to the local use of information, is long overdue. It should be entered into with great expectations--that there is value in the compiled data and that it must merely be tapped to show this value, and that if someone does not use this capability soon it may atrophy and be lost. A strong program to encourage such use, and at the same time getting some measure of its value, is in order.

This demonstration program will be concerned with the reporting, accumulation, and utilization of traffic safety data. The information flow activity is the first step in the long term loop which links the events of the traffic system to the ultimate corrective actions. Changes in the information flow system are made for the purpose of effecting improvements in the decisions made to improve conditions in the traffic system.

While the ultimate criterion for evaluating countermeasure programs is improved safety in the operating traffic system, an information flow demonstration program must be viewed in a somewhat more limited context. It could be considered as a separate self-contained system, the output of which directly affects the decisions made in most, if not all, of the other countermeasure programs.

The purpose of having a demonstration program in this field is to show that changes in the recording, accumulation, and analysis of traffic records will improve the usefulness of the available data, and will in fact generate more effective decisions by the many action agencies in the field.

This is consistent with the criteria for evaluating this system's effectiveness as 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 details of this demonstration program are developed with the thought that the machinery for effective use of traffic data is available, but that the full use of such machinery is only beginning to flower. This is particularly true in the use of state-compiled information for the local user. And the time seems to be ripe for a sort of explosion in information usage.

Data services might be viewed in two different dimensions: the first, one of intensity--from minimal (paltry) to maximal (plush); and the other, one of type--from a "pull" type service in which the user must request information when he wants it to a "push" service in which the central organization goes to a great deal of trouble to provide information which may be useful to the user even though he did not request it directly. These alternative modes of operation each have their own advantages, and it may be that the ultimate system should contain the better aspects of each. But for experimental purposes we have designed a plan for them to be applied more or less independently.

We have chosen a particular state which is carried through as an example in this report. But there are many states which are in a comparable position both in terms of equipment and activity. The ultimate choice of a site for this experimental program should depend as much as anything on the interest and willingness of the state to participate.

The experimental design calls for the selection of local communities for three levels of service: a control, a state-centered service, and a locally-centered service. The control areas will not be changed. State-centered services are primarily "push" services, tailor-made for the local community. Locally-centered services, by contrast, are "pull" activities, and the local community is basically responsible for querying the system to get information it desires.

The state-centered program is expected to have a staff of analysts and traveling salesmen whose task it is to provide a variety of assistance to local communities. They should be intimately familiar with the available data, and with the capabilities of the computational facilities. And the salemen, who in fact are giving away their product, should serve in a liason capacity to understand the local problems and match the central capability with them. We estimate that a state central analysis staff of six would be necessary in the state exemplified, and that this staff, with support, will cost between $120,000 and $200,000 per year.

The third level of service--the locally oriented one--will be really plush. There will be developed a local data base and analysis capability, although it may use the state computational facilities. There should be a close interaction with all state traffic files, but the basic work would be done by local level people through on-line computer terminals and conversational programs. We have estimated that this program (involving four cities) should cost between $400,000 and $500,000 per year.

3.3 ROAD USER PREPARATION DEMONSTRATION PROGRAM

The Road User Preparation category encompasses those programs concerned with raising the performance level of vehicle operators, passengers and pedestrians through education and training. In terms of the Highway Safety Program Standards, this area covers everything related to instructing and measuring the performance of individual road users in the motorcycle, driver education, driver licensing and pedestrian standards.

The major types of safety programs which carry out these functions may be classified as shown in Table IV.

TABLE IV. THE PHASES AND METHODS OF ROAD USER PREPARATION

	Instruction	Testing/Examining
Initiation	Child pedestrian/bicycle programs Driver education	Driver license examining
Improvement	Driver improvement Safety propaganda campaigns	Driver license re-examining

We have organized a selection from all possible countermeasures into eight groups (see Table V). Each of these deserve attention, as do some of the very large number of possibilities for combining two or more of these groups in a single program. We have identified 21 combinations which we consider to be the most useful for experimental validation, and these are noted in Figure 7 together with their relative priorities. These priorities were derived subjectively, considering the state of the art and the information federal and state planners most need concerning program effectiveness.

First priority has been assigned to three combinations (B+C+G, D+E+F+G, and B+C+D+E+F+G). We suggest that these three be implemented in a relatively long-term and fully funded demonstration program.

Similar matching fund programs might well be encouraged which will have added instrumentation and detailed experimental design attuned to the needs of NHTSA for information. Areas of lower priority may be treated in the same manner. From an analytical point of view, the value of conducting both the multiple and the singular experiments is that it enables one to sort the effects of interaction; some evidence of this would be useful for operational planners. We cannot judge the value of such choices quantitatively but together they address most of the key issues of program effectiveness.

The program we are suggesting is a comprehensive effort in driver preparation--driver education, driver license examining and driver improvement. We have considered in detail the interdependence of these three facets of driver preparation and the need for their coordination has become apparent to us in the course of this investigation. Despite the experimental difficulties associated with combining these areas into one demonstration program, we consider that this approach best meets the objectives of SCOPE--namely to demonstrate the "best we know" in traffic safety programs.

We view the purpose of a road user preparation demonstration program as one of providing information which will convince authorities to adopt useful programs in the federal Standard areas of driver education, licensing and improvement. The examination processes have generally been thought of, even in the federal Standards, as a screening device "to keep bad drivers off the road." And these processes have been largely kept separate from educational programs. There seems to

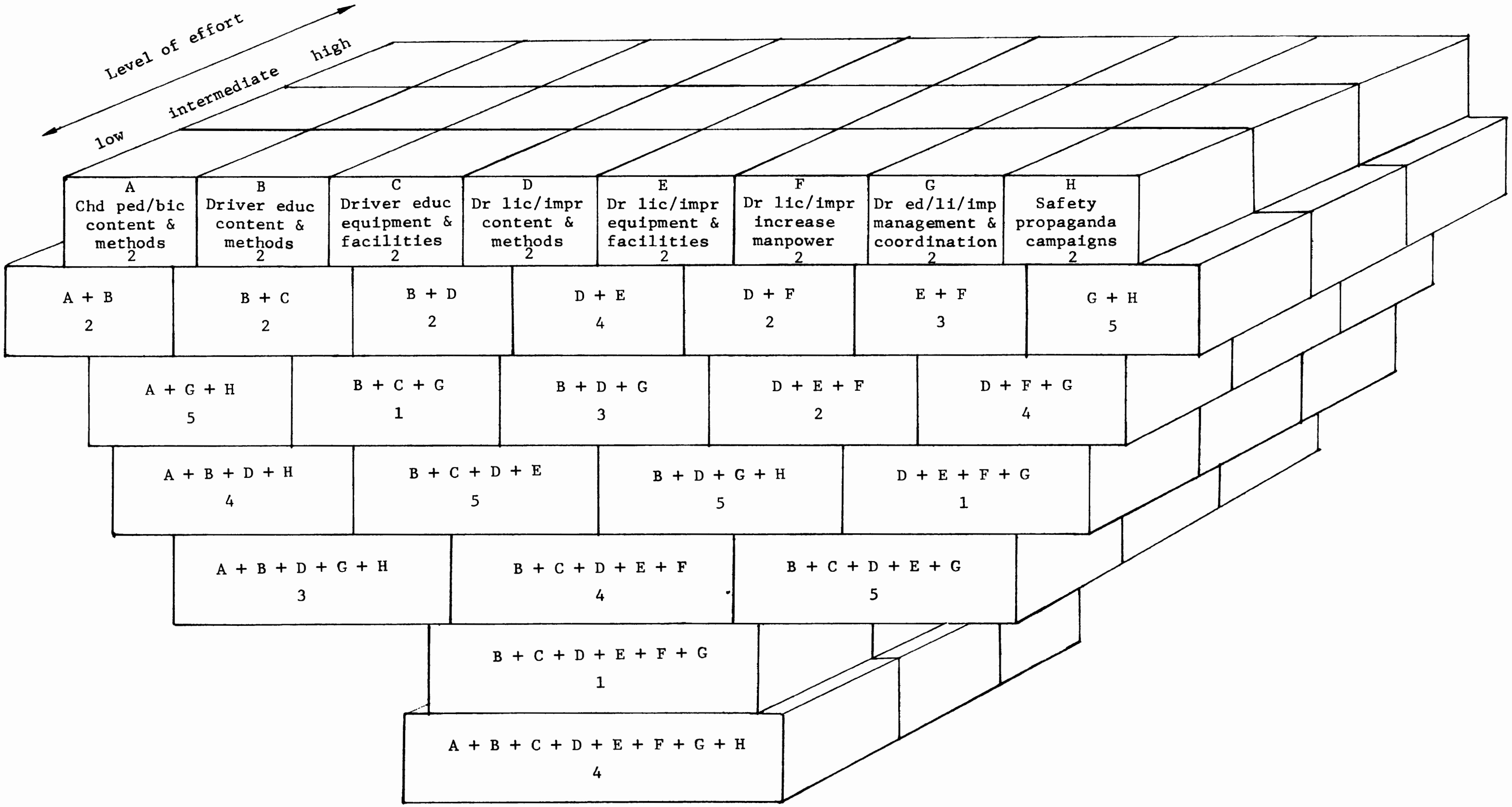


FIGURE 7. DIAGRAM OF SELECTED COMBINATIONS OF COUNTERMEASURE GROUPS (numbers denote priority for implementation)

TABLE V. COUNTERMEASURE GROUPINGS FOR ROAD USER PREPARATION

A. Child Pedestrian and 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 is to include familiarization with driver education developments, and is 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 Information(General Effort)
To make a comprehensive effort to assist those who control and practice the art of the dissemination of public information on all types in highway traffic safety.

be no doubt among state legislators of the necessity for licensing drivers, but they often do not agree on what detailed procedures the licensing authorities should use. Driver education, on the other hand, does not have unanimous support in all state legislatures; and where they do agree that it is useful, they may still argue over content. There are many innovations currently proposed in the fields of education and licensing. It is the purpose of this study to discover how some of them may best be integrated into a demonstration program together with an evaluation strategy that is meaningful to decision makers.

Evaluation in this area has proved extremely difficult, and there is a problem with reconciling the realistic needs of planners for urgent information with the fact that accurate measurement of this type of program requires careful and rather lengthy experimentation. For example, at a recent driver education conference, a prolonged discussion had taken place concerning the philosophy and scope of evaluation in this field. Suddenly a school teacher who had been patiently listening to the harangue for 2½ hours came forth with this statement: "I have been sent here by my Board of Education to find out how we can show why our funds should not be cut off next month." And it was evident from the ensuing confusion of the meeting that his question could not be answered.

Evaluation techniques suitable for driver preparation fall into three main categories:

(1) Program evaluation: primarily the auditing of what goes into the program, including the characteristics of the practitioners who execute it, and the biographical posture of road users affected; it also involves certain efficiency measures (e.g., the public's acceptance of program content and methods).

(2) Evaluation of individual proficiency against defined objectives (including instructional objectives); i.e., pencil and paper tests, skill tests, road tests, etc.

(3) Evaluation of individual "real world" performance, through unobtrusive direct observation (e.g., TV monitors, surveillance from following vehicle, etc.), and the judicious use of accident and violation records.

We do not yet have an adequate set of instruments, especially of types 2 and 3. A series of federal contracts, culminated by Harman et al. (8), led to the recommendation of a long-term plan to identify the behaviors essential to the driving task, assess their criticality, and develop instructional objectives together with instruments of types 1, 2, and 3 above. The present contract program of the NHTSA substantially reflects these recommendations, and should eventually provide the input we need not only for driver education, but also for improvement and licensing countermeasures. The completion of this work is much to be desired. But until it is completed, practitioners (such as the teacher quoted above) and planners involved in driver programs will continue to be overwhelmed by the complexity of evaluation.

In Volume III of this report we have discussed in more detail the countermeasure priorities in road user preparation as a whole. This is given as a rationale for the countermeasure groups and combinations shown in Table V. Three specific program outlines are presented as being appropriate to SCOPE.

We recommend countermeasure activity in child pedestrian and bicycle programs, and in safety information campaigns. For the child pedestrian and bicycle programs we suggest that following an effort to bring together some of the recent developments in learning activities, the teachers from a limited number of nursery and elementary schools should be trained and equipped to implement these; an observational method is suggested to measure the effect of the teaching on behavior. For information campaigns, we recommend that a small organizational structure be created in demonstration communities to coordinate media in campaigns directed towards changing driver behavior under very clearly defined circumstances (such as at specific highway locations). Observational measurement is again suggested.

We recommend a comprehensive demonstration project designed to apply the benefits (especially testing and evaluation techniques) of the NHTSA driver performance research contract program to the whole area of driver preparation. This is a three-year project involving three matched pairs of "catchment areas" of driver licensing stations; each pair contains one catchment area in a rural setting, and one in an urban environment. During the first year, certain baseline data are to be gathered in all three pairs, and a considerable amount of effort should be devoted to preparing for important changes to the driver education and the licensing and improvement functions. In the second year, these changes should take effect: the "best we know" driver education program, as defined by the current curriculum development research efforts implemented in all of the high schools, in the first pair of catchment areas; in the second pair, the "best we know" driver licensing and improvement program (the final details of which are again dependent on current research); in the third pair, both of these changes are implemented simultaneously; in all sites there are varying amounts of effort at the management level to coordinate the activities of driver education and licensing and improvement. In the third year, no further changes are made in any of the sites, but the resulting combinations of old and new education and licensing and improvement are measured carefully. The experimental design calls for comparisons between the matched sites with different amounts of change, and between various randomly selected groups of drivers within the sites. Two principal types of measurement are recommended: evaluation of the quality of coordination between the education and licensing and improvement functions, including public response; and individual driver performance measurement using a road test. The use of accident and violation data is supported only with major qualifications on its usefulness.

Finally we recognize that there will continue to be a need for advice to practitioners who are forced to make short-term evaluation in order to defend present driver preparation programs. To this end, we suggest (a) an approach for deciding what kind of evaluation is most appropriate and (b) as an interim measure, a simple driver proficiency test model that will be more closely related (albeit intuitively) to accident and violation reduction than are the traditional tests of knowledge, attitude, and skills.

3.4 VEHICLE REGULATION DEMONSTRATION PROGRAM

The principal countermeasure in the area of vehicle regulation is motor vehicle inspection; some method for the state to assure itself that the vehicles traveling on its highways are in a proper and safe operating condition. The current program of NHTSA involves efforts to determine the relative effectiveness of periodic inspection systems in state run or privately operated garages. This program might well provide some guidance as to choice within the periodic motor vehicle regime.

A number of alternative approaches to full periodic inspection exist, however, and we believe that it would be useful (in the spirit of providing convincing arguments for compliance with the Standard) to consider the alternatives along with the improved Standards as currently planned by NHTSA. A research project in which a number of programs would operate in parallel is expected to provide useful results if such factors as vehicle age, owner characteristics, mileage exposure, and environment are controlled.

Several alternatives exist for such a series of programs. These include (1) do nothing, (2) a limited voluntary inspection program such as has been proposed in Wisconsin with some rigorous random auditing of individual performance,(3) a random check lane program perhaps with two levels of intensity and/or probability of being checked, (4) a low intensity periodic program for all vehicles in a jurisdiction, (5) a high intensity periodic program, and (6) the full program as currently outlined by NHTSA.

We propose that three experimental programs which span the range of alternatives be carried out. They involve 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 I, 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. The 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 of the experimental program 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 system will operate with the final year emphasizing evaluation and recommendations. A minimum of two years of inspection operation is needed to separate impact effects from permanent effects. The experiments could occur in three states, preferably ones currently not having a standard program. More desirably a single state might adopt the three levels in different areas; the minimum level will be state-wide and the higher levels will be 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 which is usually unavailable to the states, owner maintenance practice, and vehicle inspection. Direct measurement of a program's 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 through 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 a 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 influence strongly the program's acceptance.

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

In Volume III of this report detailed attention is given to the problems of implementation and to the evaluation techniques to be used. Some discussion of operating costs and procedures is also given.

3.5 SYSTEM RESTORATION DEMONSTRATION PROGRAM

There have been a number of demonstration programs funded during the past four years in the field of emergency medical services. Some have been centered in several large cities (New York, Detroit, Miami, Philadelphia) and several have covered large open areas (Nebraska, Arizona, and Mississippi). We believe that the city demonstrations have generally shown that even a mediocre service, in terms of training and/or equipment capabilities, in a large city is likely to be much better than any rural service because of the short service times, and because of the relative frequency of emergencies and consequent greater experience of the personnel.

Statewide programs have been concerned both with communications problems and the use of helicopters. They have demonstrated, at least by example, the utility of airborne transportation techniques.

Often the basic governmental unit which needs an ambulance service and must make the decision to pay for it is the county. Further it is the in-between unit, too small for a helicopter but too large to be adequately served by a centralized service as might be appropriate for a city. We have chosen to plan an experiment at the county level with the expectation that the decision to upgrade emergency services may be made many times by county authorities over the next several years. It is these decision-makers who must decide whether to subsidize, how much tax to levy, and what kind of emergency system to support.

We have defined "system restoration" as those activities incident to restoring order to the scene of an accident--essentially the countermeasures resulting from the federal debris removal and emergency medical Standards. There is a wide range of service capability which might be provided in these areas--perhaps reaching from a volunteer with little training to a full ambulance service with highly trained medical assistants or interns and a highly coordinated debris removal team with modern equipment and trained personnel. The NHTSA Standard has suggested at least two levels of performance--the lower as a minimum to be required of all jurisdictions, and the other as a desirable goal. We have added a third in this proposed experimental program which would provide services even beyond those envisioned by the current Standard.

As in many social contexts the greatest improvement might be expected by adding some ambulance service (no matter how slight) to a community which had none. We propose, however, to begin with a control community which has a marginally acceptable service, and to proceed upwards from there. We will look primarily at some of the internal variables of the system, specifically time (time to the scene, and from the scene to the hospital as well as total service time), quality of medical assistance (to be measured in as objective a manner as possible), and diagnostic abilities of the attendants. In the field of debris removal an evaluation will consist primarily of judging case studies with some objective data regarding time and personnel abilities.

The restoration experiments are expected to take three to four years to complete; the first year for training and initial measurement, the second year to develop a smooth operation, the third year for full operation and measurement. The program should be conducted in an area of moderate population density--we have recommended a county with an area of 1000 square miles and a population of about 200,000. This is expected to provide enough data in one year of operation to assess any important changes in time and treatment.

Measures of morbidity and mortality are, as frequency counts, not likely to be very useful. This is because there can be such variation in the types of injuries, the age, sex, health of the injured, and in the kinds of vehicles involved in accidents that the expected numbers of serious injuries and fatalities will not be large enough to provide statistical validity. Nevertheless, we propose that account can be taken of these factors by preparing a number of case studies and having qualified medical personnel determine the relative value of reduced time and proper treatment in addition to the success or failure of the system on a case by case basis.

Program costs will vary with several degrees of implementation. It is estimated that a full ambulance and debris removal system, fully paid by the sponsor, would cost more than $300,000 per site-year. But it seems likely that both local and/or "402" funds might pay much of the basic program costs. Evaluation costs would not be great since much of the data taking can be done by operational personnel. It is recommended, however, that each site employ a qualified and interested physician on a part-time basis to assist in the evaluation.

There is little question in most communities of the need for both an emergency medical service and some capability for the relatively rapid removal of debris from the highway. Nevertheless, there is often a considerable problem in deciding how much of those services to buy. One might think that each community would buy what it can afford, and in a sense this is true. But there are competitive programs looking for dollars, and there is a real need for information which will assist the decision-makers (city councils, etc.) in their allocation of funds to this field. What is sought in this particular demonstration program is a relative evaluation of several levels of emergency service--ranging from the "Ford" to the "Cadillac" in quality--in terms that the local decision-makers can use to judge their worth. The post-experimental combining of the measurements into a subjective evaluation of the relative value of time and quality of service is expected to yield the desired results.

3.6 HIGHWAY REGULATION DEMONSTRATION PROGRAM

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 (i.e., crosswalks, control lights, etc.).

The formulation of an experimental countermeasure evaluation program in the highway regulation category is affected by factors such as (1) the current state of knowledge in countermeasure development; (2) the current state of countermeasure deployment; (3) practitioner's acceptance of the latest research findings; (4) the quality of existing evaluation procedures; and (5) the diversity of countermeasures that are applicable. From a review of the regulation field, it appears that the current state of countermeasure development is relatively good. Similarly, the evaluation of existing techniques is generally far better than in most other program categories. Moreover, this evaluation information is rapidly disseminated throughout the safety community by means of a variety of journals and publications of high scientific stature.

The diversity of specific countermeasures applicable to the highway category is a practical problem in experiment design. For example, the use of roadway lighting, skid resistant pavements, and traffic regulation at construction sites is called for by the highway design Standard. In terms of the interactive grouping plan discussed in detail in Volume II of this report, the experimental data matrix for such a program would consist of a number of individual countermeasure groups (lighting, pavement, etc.) and very few interactive groups. That is, the efficacy of skid resistant pavement and better lighting is likely to be additive; it is hard to imagine that better lighting will improve the skid resistance of pavements or that better traction will help the driver see better. As a further detracting factor, it is highly possible that a single site location to carry out a sizeable number of these single level groups simultaneously would be difficult to find. As a result, the demonstration programs in this category would tend to be rather small, fragmented efforts.

The real problem in the highway regulation category appears to be the current state of countermeasure deployment. Many worthwhile countermeasures such as breakaway sign supports and properly installed barrier railings have been demonstrated to be valuable means of reducing accidents, but it seems to take an inordinate amount of time to get these concepts into widespread usage. This applies to new construction practice as well as to the more expensive retrofitting of existing facilities. This type of attitude, when sufficiently prevalent, puts a severe damper on the efficacy of a demonstration program intended to show the value of certain techniques to prospective users. In effect, the people are already convinced that the measures are worthwhile; the machinery of government has simply failed to act on the conviction.

From an evaluation of all the factors presented above, we have concluded that a full countermeasure demonstration in the highway regulation category is not feasible at the present time. It seems that what needs to be done simply cannot be accomplished in a single, large-scale, demonstration program. However, the program planning methodology presented in Volume II is still felt to be applicable. Instead of serving as a guideline for the demonstration program, however, 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 efforts currently existing in this area be encouraged to combine under this plan to provide a needed degree of cohesion in the overall activities.

Nevertheless, the design of a beltway signing experiment has been described in some detail to further illustrate the evaluation concepts that we hope to promote. This topic is timely since much current effort has gone into the determination of symbolic sign effectiveness as a replacement to the conventioal legend signing. A realizable program goal for such an effort is consequently the determination of the relative efficacy of symbolic and legend signs in reducing accidents and accident potential at troublesome beltway exits.

The basic program plan employs a multi-exit beltway surrounding a large city to determine the relative efficacy of symbolic signing in informing drivers of critical information. On a given route, a number of troublesome exits can be chosen for treatment while the remaining sites are left in their pre-experiment state so as to obtain a self-controlled experiment. That is, the exits not used for sign modification can be instrumented to permit before, during, and after measurements of the same parameters measured at the actual test exits. The use of the experimental site itself as a control group is very desirable since certain characteristics of the actual population under test are evaluated over a long period of time to provide an indication of experiment stability. For instance, local strikes, tax increases, and other factors of this nature can strongly affect the attitude of the populace and hence their driving habits. Control site measurements on the actual population would provide an indication of the importance of such effects.

Because the use of highway signing is a countermeasure closely related to the accident process, it may be possible to evaluate its efficacy in terms of ultimate criteria (accident reduction). Thus, the use of accident measures will form an important part of the evaluation effort. We have suggested a number of modifications to the normal accident reporting procedure to provide more accurate data that is responsive to the problem of determining the incidence of sign-induced accidents. Before and after statistical determinations of accident rate will be used to determine if the signing modification produced a significant reduction of the observed accident rate. Exposure measures are also suggested as a valuable means of defining the population of drivers utilizing the expressway.

A change in signing produces a change in the response of the driver. That is, when the sign is observed, the driver reacts to the message, producing, in turn, a modification of the vehicle's trajectory. A direct measure of effectiveness can be obtained by observing vehicle interactions on the highway by counting specific conflict situations.* Before and after determinations of the conflict rate at each intersection under instrumentation can be used to determine if a significant (again in a statistical sense) reduction has occurred. An analysis of the accident and conflict data for the beltway may be used to refine the conflict measure and to help define the relationship between conflicts and accidents--an important relationship for future experiments.

Finally the effectiveness of the signing should be determined by the use of simulation techniques. That is, photographs of the actual signs used on the beltway should be employed with test subjects to determine such factors as glance readibility, comprehension, etc.

In summary, a careful analysis of all the data suggested in Volume III of this report should permit a determination of the relative efficacy of symbolic and legend signing in reducing accidents and accident potential. Moreover, this knowledge will be backed by the why of the countermeasure operation (i.e., why did the signing produce the results that it did) so that the experience gained in this experiment may be extrapolated to other sites.

---

*See for example, S. R. Perkins, GMR Traffic Conflicts Technique Procedures Manual, Research Publication No. GMR-895, August 11, 1969

4. CONCLUSIONS

The fact that there is both a real and a felt need for more information regarding the value of various social programs aimed at accident prevention is evident: both legislative bodies and administrative units continue to ask for proof of either the absolute or relative value of accident countermeasures.

While there have been numerous attempts to derive such information in the form of cost-benefit ratios, these have been rather unsuccessful except where system changes relate directly to the accident or injury process (i.e., restraint systems or guard rails). It is useful, however, to demonstrate that system changes have a positive effect as measured at some point other than the ultimate measure of injury or fatality. With this in mind, we have suggested intermediate measures of effectiveness related directly enough to the change to be identified with it, but also related well enough to the accident probability to be considered responsible for the reduction in accidents or injuries.

Specific experimental designs have been prepared based on this concept that the measurement of an intermediate effect of a change is most useful. We have recommended that a program to make such measures be undertaken, with the expectation that the measurements will be directly useful to decision-makers. While the many suggested countermeasures derive initially from the current set of sixteen federal Highway Safety Program Standards, we have grouped these into only six programs for the purpose of experimental design.

It is clear that there will continue to be operational and budgetary decisions made in the field of highway safety, and that factual information about the effectiveness of the many possible countermeasures will be sought. We believe that the programs outlined in this document will lead logically to better decisions and thence to a higher degree of safety on the country's highways.

REFERENCES

1. Havelock, Ronald G. and Markowitz, Elizabeth A., A National Problem-Solving System: Highway Safety Researchers and Decision Makers, Final Report No. FH-11-6900, Center for Research on Utilization of Scientific Knowledge, Institute for Social Research, The University of Michigan, Ann Arbor, May 1971.

2. Herms, Bruce F., Pedestrian Crosswalk Study: Accidents in Painted and Unpainted Crosswalks, California 402 Project, PS 69-2-001, San Diego, August 1970.

3. Womack, J.C. Median Barriers and Accident Prevention, California State Highway Department, Sacramento, October 10, 1966.

4. Center Barriers Save Lives, New Jersey Department of Transportation, New Brunswick, August 1968.

5. Van Vechten, C. Thomas, Statistical Analysis of Accident Data, 2d ed., Government of the District of Columbia, Department of Highways and Traffic, Bureau of Traffic Engineering and Operations, Washington, D.C., May 1969.

6. Campbell, B.J., Personal Communication, 1971.

7. Wilson, E.B., An Introduction to Scientific Research, McGraw-Hill, 1952.

8. Harman, H.H., et al., Evaluation of Driver Education and Training, Highway Research Board, Washington, D.C., March 1969.