Report No. FHWA-TO-80-1

Evaluation of Traffic Operations, Safety, and Positive Guidance Projects

OCTOBER 1980

VOLUME II IN A SERIES ON POSITIVE GUIDANCE

POSITIVE GUIDANCE

U.S. Department of Transportation
Federal Highway Administration
Office of Traffic Operations

NOTICE

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof.

The contents of this report reflect the views of the Office of Traffic Operations, Federal Highway Administration, which is responsible for the facts and accuracy of the information presented therein. The contents do not necessarily reflect the official views or policy of the Department of Transportation. This report does not constitute a standard, specification, or regulation.

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA-TO-80-1		

4. Title and Subtitle Evaluation of Traffic Operations, Safety and Positive Guidance Projects	5. Report Date October 1980 6. Performing Organization Code
7. Author(s) Harold Lunenfeld	8. Performing Organization Report No.
9. Performing Organization Name and Address Human Factors Branch Office of Traffic Operations Federal Highway Administration Washington, D.C. 20590	10. Work Unit No. (TRAIS) 11. Contract or Grant No.
12. Sponsoring Agency Name and Address Traffic Performance and Programs Division Office of Traffic Operations Federal Highway Administration Washington, D.C. 20590	13. Type of Report and Period Covered 14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

This volume is for use by highway and traffic engineers and technicians in implementing an evaluation of Traffic Operations, Safety and Positive Guidance projects. Evaluation should begin in the planning stage where evaluation designs are selected and measures of effectiveness (MOE's) identified. It should continue during the solution development phase where the measures are used as diagonostics and culminate in an evaluation where differences are in the measures taken before and after the improvement are used to assess the effectiveness of the solution. The Evaluation report provides a "cookbook" approach to implementing the procedure. This enables those engineers and technicians who are not versed in accident or accident surrogate MOE's, experimental design or statistical analysis to determine whether their solution is effective. Each step in the evaluation is structured in terms of inputs, outputs, and the logic involved in its execution. Tables and worksheets lead to the development of a detailed evaluation plan, data collection procedure and data analysis routine. Among the factors detailed are the selection of appropriate MOE's, overcoming threats to the validity of the evaluation, ways to assure "Before" and "After" comparability, how to select the proper statistical test, establishing an appropriate confidence level and recognizing the importance of practical significance.

17. Key Words	18. Distribution Statement
Evaluation, traffic operations, safety projects, spot improvements, Positive Guidance, experimental design, measures of effectiveness, statistical analysis, practical significance.	No restriction. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22151

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	224	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

26136

JUN 27 2000

FOREWORD

As an engineering psychologist trained in research, experimentation and statistics, I assumed that it would be relatively easy to prepare a project evaluation treatise. However, when I attempted to write a brief evaluation section for the Users' Guide to Positive Guidance, I found that most of my intended audience, engineers and technicians in the "real world," did not have the experience and training to conduct a valid evaluation. I also found that there was no single reference that covered all aspects of evaluation. Finally, I found that I could not readily define all aspects of project evaluation.

This led to the effort that resulted in this report. From the beginning, I concluded that the only way to develop an evaluation procedure for operations personnel with limited backgrounds was to put myself in their shoes. I followed a human factors approach by determining and defining the evaluation process in terms of inputs, outputs and flow charts. I surveyed State and local jurisdictions to identify available resources--personnel, equipment, reference material, etc. I identified who would be conducting the evaluation and determined the time and effort that they could devote. Finally, I considered what I would need to perform an evaluation, given background, time and resource limitations.

I concluded that the only practical approach, given the range of projects, objectives, needs and available resources, was to develop a step-by-step procedure in the form of a self contained "cookbook". Thus, the brief evaluation section I began writing evolved into this report. I was concerned that the document's size would discourage individuals from using the procedure. However, I concluded that most of the report is actually figures, tables, forms and examples; that a large range of situations had to be covered; that users of the procedure would not have to go through all the material; and, that the report would be used as a resource document. I thus found that it would not be possible to delete material and still achieve the objectives of the report. I also concluded that the time spent in following the procedure would pay dividends in time and effort saved in the field, and in the utility of the conclusions derived from the conduct of a suitable evaluation.

I am indebted to a large number of individuals in the Federal Highway Administration, in State and local jurisdictions, and in private industry for their help in preparing and reviewing this report.

I wish to single out Dr. Wallace G. Berger of the U.S. Senate Staff, Mr. Robert S. Hostetter of the Institute for Research, and Mr. Theodore J. Post of BioTechnology, Inc., for their assistance and inputs.

Finally, I am most indebted to Mr. Gerson J. Alexander, Chief, Human Factors Branch, Office of Traffic Operations, Federal Highway Administration, for his suggestions, support and encouragement.

Washington, D.C. Harold Lunenfeld

TABLE OF CONTENTS

	Page
Foreword	i
Executive Summary	iv
Introduction	1-1
Scope	1-1
Background	1-3
Overview	1-4
The Evaluation Plan	2-1
Overview	2-1
Evaluation Plan Factors	2-4
Step 1. Identify Project Objectives	2-7
Step 2. Specify "Before" Conditions	2-11
Step 3. Select Measures of Effectiveness	2-15
Step 4. Select Measurement Technique and Operationally Define MOE's	2-21
Step 5. Specify Implementation and Acclimation Period	2-30
Step 6. Select Evaluation Design	2-35
Step 7. Specify Statistical Tests	2-40
Step 8. Specify Confidence Level	2-46
Step 9. Develop Sampling Plan	2-50
Step 10. Prepare Data Collection Plan and Schedule	2-55
Collection of Evaluation Data	3-1
Overview	3-1
Step 11. Perform Pre-Data Collection ("Before" Phase)	3-3
Step 12. Collect "Before" Data	3-6
Step 13. Reduce and Summarize "Before" Data	3-8
Step 14. Schedule Improvement Application	3-28
Step 15. Allow for Acclimation	3-29
Step 16. Perform Pre-Data Collection ("After" Phase)	3-32
Step 17. Collect "After" Data	3-34
Step 18. Reduce and Summarize "After" Data	3-36

	Page
Assessment of Results	4-1
Overview	4-1
Step 19. Summarize and Compare Traffic and System Performance MOE's.	4-3
Step 19A. Summarize and Compare Accident MOE's	4-8
Step 20. Determine Statistical Significance	4-14
Step 20A. Poisson Distribution Test	4-16
Step 20B. t Test.	4-20
Step 20C. F Test.	4-24
Step 20D. Z Test.	4-28
Step 20E. Chi Square Test	4-30
Step 21. Determine Practical Significance.	4-34
Step 22. Report Findings and Recommendations	4-37
Glossary of Terms	5-1
References and Bibliography.	6-1
Appendix A. Statistical Factors	A-1
Appendix B Procedure to Summarize and Compare Accident MOE's	B-1
Appendix C. Forms.	C-1

EXECUTIVE SUMMARY

With current emphasis on safety and the efficient operation of existing facilities, the need to implement projects that provide optimum traffic control and driver information improvements has assumed increased importance. This has resulted in the requirement for engineers, technicians, and other design and operations personnel to evaluate their improvements to insure that they are effective.

Evaluation should be an integral part of traffic operations, safety and Positive Guidance projects. The process should begin in the planning stage where an evaluation design is selected and measures of effectiveness (MOE's) identified, continue during improvement development where the measures are used as diagnostics, and culminate in the post-implementation phase where differences in the MOE's are used to assess effectiveness.

Since most State and local jurisdictions possess neither the expertise nor are able to hire specialists familiar with evaluation designs, statistics or surrogate measures, a procedure is needed to enable personnel to perform a suitable analysis and evaluation. Such a procedure has been developed in conjunction with the Positive Guidance program. This report describes the procedure and shows its application. The procedure provides a step-by-step "cookbook" approach to evaluation. Each step is structured to identify inputs and outputs and show the process. Tables, worksheets, flow charts and examples lead to the development of a detailed evaluation plan, a suitable data collection procedure, and an appropriate data assessment routine.

The first phase, the Evaluation Plan Development, takes into account factors that affect the suitability of the evaluation and the validity of the results. These factors include: Appropriateness; validity; reliability; regression toward the mean; representativeness; novelty effects; changes over time; comparability; experimental designs; statistical factors; randomization; obtrusiveness; and stability.

The following ten steps are performed to develop the Evaluation Plan:

STEP 1. Identify Objectives - In this step, a project is identified as either Safety and Hazard Potential Reduction or Traffic Operations and Flow Optimization. Objectives are then formulated in terms of accident reduction, traffic performance improvements and system performance improvements.

STEP 2. Specify "Before" Conditions - In Step 2, those climatological, environmental and traffic stream conditions representative of the problem are specified.

STEP 3. Select Measures of Effectiveness (MOE's) - During the third step, the objectives are translated into measures used to diagnose the problem and evaluate the improvement.

STEP 4. Select Measurement Techniques and Operationally Define MOE's -Data collection methods are selected and the operations or procedures employed in distinguishing the selected MOE from other measures are defined in precise, observable terms in Step 4.

STEP 5. Specify Implementation and Acclimation Period - In Step 5, the length of time between the collection of "Before" and "After" data required to develop and apply an improvement and for drivers to acclimate to it is specified.

STEP 6. Select Evaluation Design - Either of the two recommended designs, the "Before-After" or the "Before-After with a Control Site" is selected in this step.

STEP 7. Specify Statistical Tests - In the seventh step, five statistical tests; Poisson, t, F, Z, and Chi Square are recommended and criteria provided for their use.

STEP 8. Specify the Confidence Level - The concept of confidence (level of significance) used to express the probability of the results being due to chance is described. Since statistical tests are evaluated in terms of confidence, a suitable level, ranging from 80% (.20) to 95% (.05) is specified.

STEP 9. Develop Sampling Plan - The minimum sample size and the sampling plan used are developed in Step 9.

STEP 10. Prepare Data Collection Plan and Schedule -During the final step in this phase, all aspects of the Evaluation Plan are brought together in a Data Collection Plan and Schedule.

The Collection of Evaluation Data phase follows the development of the Evaluation Plan. The plan is put into effect in the first part of the phase when "Before" data are collected. Following the development of the improvement, the second part of the phase is involved with scheduling the improvement and allowing for acclimation. Following a suitable acclimation, the "After" data are collected.

This phase consists of the following eight steps:

<u>STEP 11.</u> <u>Perform Pre-Data Collection Activites ("Before" Phase)</u> - During the initial step, logistical matters that have to be completed prior to collection of "Before" data are considered. Any needed pilot testing is conducted.

<u>STEP 12.</u> <u>Collect "Before" Data</u> - The required "Before" data are collected during this step.

<u>STEP 13.</u> <u>Reduce and Summarize "Before" Data</u> - Appropriate descriptive statistics are applied to the "Before" data in Step 13.

<u>STEP 14.</u> <u>Schedule Improvement Application</u> - The "Before" phase is assessed and the improvement is scheduled in Step 14.

<u>STEP 15.</u> <u>Allow for Acclimation</u> - In this step, an appropriate time period following the application of the improvement is allowed to elapse.

<u>STEP 16.</u> <u>Perform Pre-Data Collection Activities ("After" Phase)</u> - Step 16 is a repeat of Step 11 for the "After" phase.

<u>STEP 17.</u> <u>Collect "After" Data</u> - The required "After" data are collected in Step 17.

<u>STEP 18.</u> <u>Reduce and Summarize "After" Data</u> - The same descriptive statistics used in Step 13 are applied to the "After" data in Step 18.

The final phase assesses the results, thereby determining whether and to what extent project objectives have been met.

There are four steps in the Results Assessment phase.

<u>STEP 19.</u> <u>Summarize and Compare MOE's</u> - Data from Steps 12 and 18 are summarized and compared in Step 19.

STEP 20. Determine Statistical Significance - The appropriate statistical test is applied to the results at the preselected confidence level to determine whether the improvement is statistically significant.

STEP 21. Determine Practical Significance - In this step, the results are examined in the context of whether or not the improvement has led to meaningful and, in some cases, economically beneficial solutions.

STEP 22. Report Findings and Recommendations - The final step is to prepare and distribute a final report indicating findings, conclusions and recommendations.

INTRODUCTION

Scope

This report is the first in a series of documents published in conjunction with the Positive Guidance program. Its purpose is to give engineers and technicians a tool to evaluate highway information system related spot and short segment improvement projects. Other documents in the series will cover the collection and use of field data, the Positive Guidance procedure, and the results of the Positive Guidance Demonstration Program.

The intent of this report is to provide the user with a step-by-step approach to project evaluation. Its specific focus is on the evaluation of traffic operations, safety, and Positive Guidance projects. To apply the procedure requires that the user have an appreciation of the range of evaluation strategies used for the various types of projects. While consider-able discussion is given to safety projects and accident measures of effectiveness (MOE's), emphasis in this report is on traffic operations and Positive Guidance projects and their associated traffic and system performance MOE's.

This document provides the user with several types of information to serve as a guide in setting up an evaluation and applying the procedure.

- Background and Rationale - Principles, concepts and terminology are given for each phase on the evaluation, and for individual steps.
- Inputs, Outputs and Flows - Each step is structured in terms of the logic involved in its execution to enable users to see how things fit together.
- Tables, Charts, Figures and Examples - Forms and examples are provided to enable personnel to progress in a stepwise manner.

Procedures Overview

The evaluation procedure consists of three phases and twenty-two steps. The Project Planning phase has ten steps, all designed to produce an Evaluation

Plan and schedule. An Improvement Development phase is comprised of eight steps required to implement the plan and collect the necessary "Before" and "After" data used to diagnose the problem and evaluate the effectiveness of the improvement. Finally, the Improvement Evaluation phase consists of the four steps used to reduce and analyze the data, test for significance and report the findings.

Document Organization

The report is organized in eight parts. Page, figure, table and exhibit numbering is consecutive within each part. Exhibits are used to provide examples of each step. The report is arranged as follows:

Executive Summary
Introduction - Part 1
The Evaluation Plan - Part 2
Collection of Evaluation Data - Part 3
Analysis of Results - Part 4
Glossary of Terms - Part 5
Reference and Bibliography - Part 6
Appendices and Forms - Appendix A - C

Evaluation Report Uses

The primary use of the evaluation report is as a working document and training tool to enable engineers and technicians to set up and conduct a highway improvement project evaluation. The report is organized to allow individuals who are not familiar with evaluation design and statistical methods to proceed in a step-by-step manner. Because the document is structured to encompass a range of project types, situations and needs, all steps in the procedure may not be required in each case. The user should tailor an evaluation to the needs of his or her particular project.

Once users have become familiar with the material in this report, they can combine or skip steps, as well as modify forms or procedures to suit their particular needs. However, material for all applications is included for completeness and to show how all parts fit together.

Background

Implicit in the application of a systematic procedure such as Positive Guidance is the assumption that proper implementation of the steps in the procedure will result in suitable and effective improvements. Evaluation is an integral part of Positive Guidance. An evaluation is recommended for all traffic operations and safety projects as well, since it is always necessary to assess whether, and to what extent, improvements have succeeded in ameliorating problems.

Evaluations achieve a number of important goals. In addition to providing data about the effectiveness of the project, they identify the need for further development and enable decisionmakers to accept or reject similar improvements in similar applications. Evaluations thus provide answers about the effectiveness of a specific project and information for future applications.

An evaluation compares safety and/or traffic operations before and after an improvement has been developed and applied to determine whether performance has been effectively improved. There are four basic methods: (1) Analytical; (2) empirical; (3) economic; and (4) subjective. The analytical method assesses the improvement prior to its being implemented by comparing products of solution development steps (such as the "expectancy analysis" and "information load profile" of Positive Guidance projects) for improvement. The empirical method, described in this report, compares differences in measures of effectiveness (MOE's) for statistical and practical significance. When differences are found to be statistically significant, an economic analysis is often performed, data permitting. Finally, when it is not feasible to perform the empirical method, the solution can be subjectively evaluated by having peers, experts, and/or ordinary drivers providing their judgment on the effectiveness of the improvement.

Application of the empirical method requires the user to understand and apply statistical procedures. Since many engineers and technicians may not be familiar with evaluation techniques, statistical methods, or certain MOE's, a step-by-step "cookbook" approach is provided. However, it is recommended that those unfamiliar with statistics and related disciplines seek assistance in applying the statistical procedures. This assistance should be obtained in the Project Planning phase, at the time that the Evaluation Plan is under development. Delaying this action until the improvement has been developed and in place may lead to wasted effort or erroneous conclusions.

Overview

Figure 1-1 shows the relationship of the evaluation to the overall project development process. Evaluation should begin during the Project Planning phase, where an Evaluation Plan is generated. During the Improvement Development phase, "Before" measures are used as diagnostics to develop improvements, and, following improvement installation and a suitable acclimation, "After" data are collected. Finally, during the Improvement Evaluation phase, a comparison is made of "After" data and "Before" data to determine whether differences are statistically significant. This determination, coupled with an assessment of practical significance and/or with an economic analysis, provides an indication of the effectiveness of the improvement.

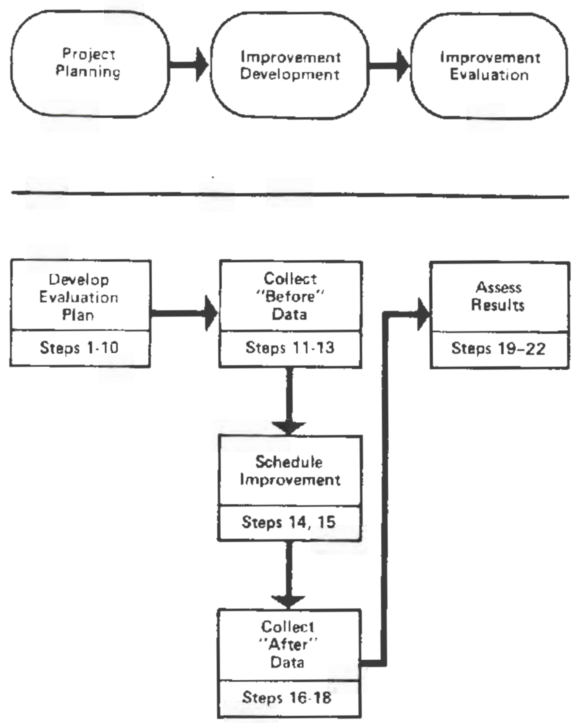


Figure 1-1.
Evaluation Overview

THE EVALUATION PLAN

The evaluation process starts at the beginning of the project with the development of a "Before" Data Collection Plan. The "Before" Plan, which serves as a diagnostic for developing the improvement, when completed by the addition of an "After" Data Collection Plan, becomes the Evaluation Plan. The time and effort spent in generating a comprehensive plan during the project planning phase assures smooth data collection, proper data analysis and meaningful results.

Much of the data used to determine the nature of the problem and to generate the improvement (i.e., diagnostic information) is also used to evaluate its effectiveness. Table 2-1 provides a list of inputs to the "Before" Data Collection Plan. The inputs come from a site survey and review of historical data that should be conducted at the onset of all projects. The reader is referred to Report FHWA-TO-80-2 for a discussion of these factors.

Overview

The purposes of the evaluation are: To determine whether the improvement was effective in reducing the problem; whether the improvement will continue to be effective in the future; and whether the improvement will be effective for similar problems. These determinations are primarily based on inferences made from a statistical assessment of differences in measures taken before and after the application of the improvement. Effectiveness is inferred on the basis of the suitability and validity of the evaluation and the attainment of significant results.

Figure 2-1 presents a functional flow of the steps that are used to develop the Evaluation Plan. The ensuing material is structured to enable project personnel to apply each step sequentially. This step-by-step approach is designed to facilitate implementation of the evaluation through the use of flow diagrams to explain each step and forms and examples to illustrate the end product of the step.

It is recommended that the user establish and maintain an evaluation file for each project. Forms filled out in each of the steps will serve as inputs to the Evaluation Plan.

TABLE 2-1. INPUTS TO THE "BEFORE" DATA COLLECTION PLAN

I. Site Survey:

- Location characteristics - geometrics, alignment, pavement
- Vantage points for data collection
- Locations where accidents occurred
- Location of traffic control devices
- Hazards
- Sources of driver confusion
- Vehicle & traffic actions (erratic maneuvers, conflicts, etc.)
- Environmental factors (weather, illumination, etc.)

II. Historical Data:

A. Traffic Data

- Volume counts
- Vehicle speeds
- Travel time and delay
- Density, gaps, spacing
- Occupancy
- Input-output
- Capacity and delay
- Signal timing

B Accident Data

- Location
- Type
- Severity
- Environmental condition
- Day and date
- Actions of drivers
- Contributory conditions

C. Other Sources

- Complaints
- Police, maintenance, etc.

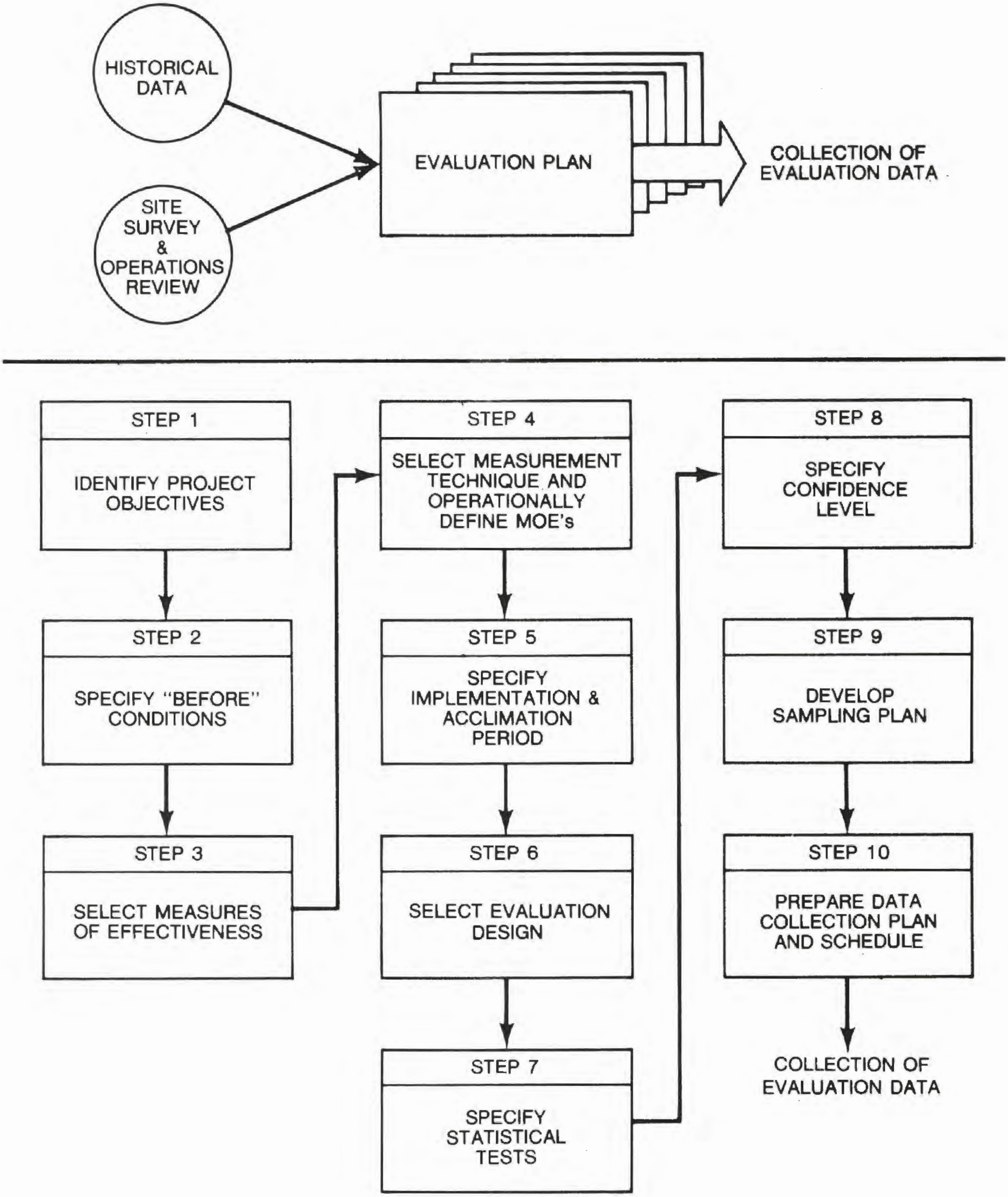


Figure 2-1.
Evaluation Plan - Functional Flow

Evaluation Plan Factors

There are a number of factors that must be taken into account by the evaluation. These factors, discussed below, affect the suitability of the comparisons and the validity of the results.

Appropriateness - The suitability of the evaluation rests on the proper questions being asked using the proper information to answer these questions. An evaluation might not focus on relevant issues unless objectives and measures are appropriate. For example, it would not be appropriate to evaluate a rural narrow bridge accident reduction project using intersection capacity MOE's. Steps 1 and 3 of the plan address this factor.

Representativeness - The conditions under which MOE's are collected may affect the extent to which inferences can be made and the results generalized to other problems. Data collected when conditions are not representative of the problem may not be valid. For example, if strangers getting lost at a freeway interchange is a major problem, then data collected during weekday commuter hours may not be as representative as data collected on weekends and holidays. Step 2 of the plan considers representative conditions and driver and vehicle populations.

Validity - The validity of the evaluation refers to the correctness of the MOE's used to diagnose the problem and the reasoning process used to assess the results. An invalid evaluation yields incorrect conclusions. For example, a centerline encroachment is not a valid MOE for delay. Similarly, an evaluation that does not take "changes over time" into account may not be valid. Step 3 is directed toward selecting valid MOE's while all steps in the plan are directed toward insuring the correctness of the reasoning process.

Reliability - A reliable measure is stable and repeatable. Measuring equipment and data collection personnel must obtain similar data each time a measurement is made. If not, the results will not be valid. For example, an uncalibrated radar speed meter or an untrained data collector using an electronic stopwatch may collect unreliable speed data. Step 3 considers reliable data collection techniques, while Step 4 accounts for repeatability and accuracy by operationally defining MOE's.

Regression Toward the Mean - When repeated measurements of any variable are not highly correlated, then unusually high or low values measured one time may be followed by values closer to average the next time. For example, if accidents at a location are modestly correlated from year to year, and an unusually high number of accidents has occured during a given year, chances are that even without applying any improvement, the number of accidents at the site will be reduced during the next year. Accidents should thus not be the sole MOE, particularly when there is an unusually high accident frequency during a given year. In Step 3, alternate MOE's to accidents are given, while Step 5 considers a suitable time period for accident frequency to average out.

Novelty Affects - Repeat drivers may respond in a transitory manner to changes in a site's information system. After a while these drivers become accustomed to the changes. For example, both strangers and repeat drivers will be affected by new guide signs, even though the target population has a high proportion of strangers. As the novelty wears off, repeat driver behavior normalizes, and the stranger's behavior can be measured. Step 5 considers the novelty affect.

Changes Over Time - It is desirable that no major changes occur during the implementation and acclimation period. However, changes are possible, particularly when the implementation and acclimation period is long. Examples include changes in traffic volume, traffic makeup, origin and destination distributions and enforcement procedures. Step 5 discusses ways that changes over time can be accounted for in performing the evaluation.

Comparability - The only thing that should change at a site between the "Before" and "After" period is the improvement. All other conditions should be comparable. For example, "After" data should not be collected at night if "Before" data were collected during the day. In Steps 9 and 10, ways to assure comparability are discussed.

Evaluation Designs - An appropriate analytical framework is required to assure that the Evaluation is valid. This is the Evaluation Design discussed in Step 6. An example of an invalid design is one where an improvement is applied without collecting "Before" data, thereby providing no basis for comparison of "After" data.

Statistics - If the statistics used to reduce and analyze the results of the evaluation are inappropriate, then it will not be possible to gauge significance. For example, merely comparing "Before" and "After results without applying an appropriate test of significance will not enable the evaluator to determine whether the results are statistically significant. Step 7 provides statistical test selection criteria and applications guidance. In Step 8, issues involved in the selection of a confidence level are discussed.

Sampling Factors - The way in which samples are drawn and the amount of data collected affect the suitability of the evaluation. An inappropriate sample may not be representative of the target population or may yield erroneous results. For example, measuring all spot speeds of vehicles in a platoon, rather than the lead vehicle, would yield unrepresentative speeds. Too small an accident sample would not allow for the use of the Poisson Distribution test. Sampling factors are discussed in Step 9.

Obtrusiveness - The method used to collect data can influence driver behavior and invalidate the results of the evaluation. If equipment and personnel are visible to drivers, their presence may cause modifications in speed and path which can be erroneously attributed to the improvement. Step 10 deals with problems associated with obtrusiveness.

Stability - The accuracy of equipment and personnel can deteriorate with time. Instability can thereby adversely affect data reliability. Ways to control for stability include calibration of equipment and training of personnel. Step 10 considers stability issues.

STEP 1. Identify Objectives

It is important to identify objectives at the beginning of the evaluation in order to formulate the measures that will be used to determine whether, and to what extent, the improvement is achieving its goals. It should be recognized that there are often several project objectives that can be identified.

Most projects are either safety projects designed to rectify a high accident location or a location with a high hazard potential, or traffic operations projects designed to rectify a problem in traffic operations or to optimize traffic flow. Objectives associated with these projects can be classified into one or more of the categories of Accident Reduction, Traffic Performance Improvements, or System Performance Improvements.

Accident Reduction - The explicit purpose of most safety projects is to reduce the total number of traffic accidents (fatal, injury, property damage) and their severity. Often, a project is undertaken to reduce specific kinds of accidents, such as sideswipes, or accidents occurring under a specific environmental condition, such as wet weather. When a specific type of accident is identified as an accident reduction objective, it should be the predominant one, based on a review of historical accident data. There should also be sufficient accident experience for total accidents as well as for specific types to test for statistical significance. Unless a large accident reduction is expected, a considerable number of "Before" accidents will be required to achieve significance. It should be recognized that 3 years of suitable "Before" accident data are required to evaluate accident reduction objectives.

Traffic Performance Improvement - Many safety projects are implemented to improve locations with a high hazard potential. These locations may not have sufficient accident experience to evaluate statistically, often because of low volume. Typical sites include narrow bridges, lane drops, highway rail grade crossings, long term construction zones, intersections and interchanges. Even when accident experience is high, the 3 year waiting period required to evaluate accident reduction may not be acceptable. In these instances, traffic performance measures serve as surrogates for accident reduction. In other cases, a high incidence of hazardous maneuvers or deviations from proper speed and path leads to project implementation. Further, traffic performance improvements are generally concomitant with accident and hazard potential reduction. For example, reducing last minute lane changes and speed variance often leads to accident reduction.

System Performance Improvements - Projects are frequently implemented to improve highway system performance. Optimizing traffic flow by increasing capacity or reducing delay is an objective of numerous projects. In addition, system performance improvements are often concomitant benefits of safety and traffic performance improvements. For example, a project designed to reduce a hazard potential at an intersection with restricted sight distance generally affects the intersection's capacity and delay characteristics.

The identification of project objectives is based on a review of information gathered during the initial phase of most projects (see Table 2-1). The List of Typical Objectives presented in Table 2-2 also serves as a guide in completing this step.

TABLE 2-2. TYPICAL PROJECT OBJECTIVES

TYPE	PROJECT	OBJECTIVES
SAFETY AND HAZARD POTENTIAL REDUCTION	ACCIDENT REDUCTION	TO REDUCE: . Total Accidents . Fatal Accidents . Injury Accidents . PDO Accidents . Dry Accidents . Wet Accidents . Snow, Ice Accidents . Overturn Accidents . Collision with Motor Vehicle, Fixed Object, Pedestrian, Animal, etc. . Two Vehicle Accidents . Angle Accidents . Head-On Accidents . Rear-end Accidents . Sideswipe Accidents . Others
SAFETY AND HAZARD POTENTIAL REDUCTION TRAFFIC OPERATIONS AND FLOW OPTIMIZATION	TRAFFIC PERFORMANCE IMPROVEMENTS	TO REDUCE: . Erratic Maneuvers . Traffic Conflicts . Speed Variance . Others TO IMPROVE: . Lateral Placement . Speed Profile . Others
TRAFFIC OPERATIONS AND FLOW OPTIMIZATION	SYSTEM PERFORMANCE IMPROVEMENTS	TO REDUCE: . Delay . Traveltime . Others TO IMPROVE: . Capacity . Level of Service . Energy Efficiency . Others

Figure 2-2 shows the Step 1 flow. Objectives should be listed on a form such as the one shown in Exhibit 2-1. In filling out the form, the evaluator should first identify the primary objective and then determine concomitant or secondary ones. It is useful also to indicate justification for selection in order to limit the listing to those objectives amenable to evaluation.

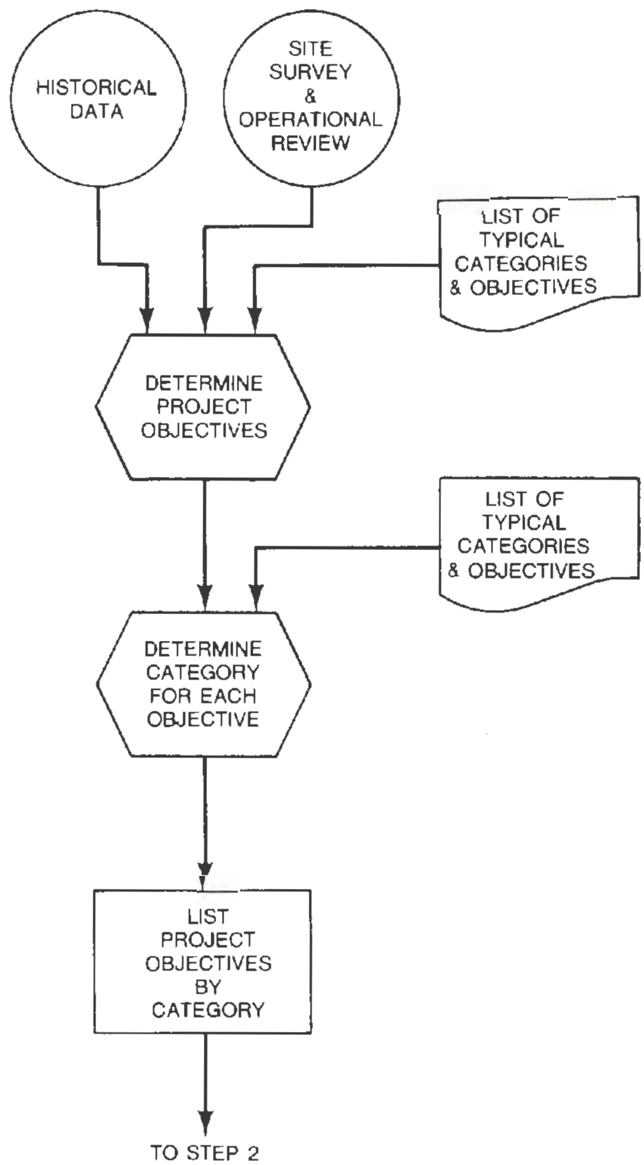


Figure 2-2.
Flow, Step 1 - Identify Objectives

OBJECTIVES LISTING—STEP 1		
PROJECT: *Intersection of 12th and Farrell Street*		
PROJECT NO.: *123B*		
EVALUATOR/DATE: *HL 7/6/79*		
CATEGORY	**OBJECTIVE**	**COMMENTS**
ACCIDENT REDUCTION	*(1) To reduce total no. of accidents* *(2) To reduce right-angle accidents*	*This is a primary objective because of 25 accidents in 1978 including 16 right-angle accidents*
TRAFFIC PERFORMANCE IMPROVEMENT	*To reduce left-turn cross traffic conflicts*	*This is a secondary objective–site survey indicated a high level of conflicts*
SYSTEM PERFORMANCE IMPROVEMENT	*To reduce delay*	*This is a primary objective–project funded to improve operations and reduce delay.*

Exhibit 2-1.
Sample Objectives Listing Form

STEP 2. Specify "Before" Conditions

The more representative the conditions under which "Before" data are collected, the more appropriate will be the evaluation. "Before" conditions are determined from an analysis of the range of environmental, driver and traffic conditions identified during the initial stages of a project. This information is used to specify target conditions and to identify target driver and/or vehicle populations.

While it is generally advisable to collect data under a representative range of conditions, it is usually neither possible nor desirable to gather data for every condition. It is often necessary to limit data collection due to time and monetary constraints. In addition, the nature of the problem may be such that only certain conditions are warranted. For example, data collected during dry conditions would be irrelevant if there is evidence that the problem being attended to only occurred under wet pavement conditions.

A review of problems and/or accidents generally focuses in on target driver and/or vehicle population as well as "Before" conditions. For example, if strangers getting lost is identified as the major problem, non-repeat drivers should be maximized. This can be accomplished by collecting data during weekends, holidays, and the tourist season. Similarly, when most accidents occur after dark, MOE's data collection should be performed at night.

The procedure to follow in specifying "Before" conditions is to develop a set of tentative "Before" conditions including target driver and vehicle populations, based on a review of historical data, a site survey and operations review, and an analysis of the problem. The information contained in Table 2-3 serves as a guide. Once target conditions are tentatively identified, a feasibility assessment should be started (see Steps 3 and 4). Such factors as availability of equipment and personnel; potential data collection methods (e.g., time lapse is usually not feasible at night); time schedules (e.g., it may not be possible to wait for a particular season); and, uniqueness of the condition (e.g., fog or blowing dust) should be considered to finalize "Before" conditions.

TABLE 2-3. TYPICAL "BEFORE" CONDITIONS

Category	Condition
Weather	Clear Rain Fog Snow Other
Illumination	Day Night Twilight Illumination (Specify)
Pavement	Dry Wet Icy Type (Specify) Condition (Specify)
Traffic Volume	AADT (Specify) Peak (Specify) Hourly (Specify) Other
Season	Spring Summer Fall Winter Tourist
Target Populations	Drivers - Locals Strangers Tourists Commuters Vehicles - Passenger Commercial Recreational Other
Others	Foliage Striping Seasonal Events Legal/Enforcement

Figure 2-3 shows the Step 2 flow. A form for specifying "Before" conditions is presented in Exhibit 2-2.

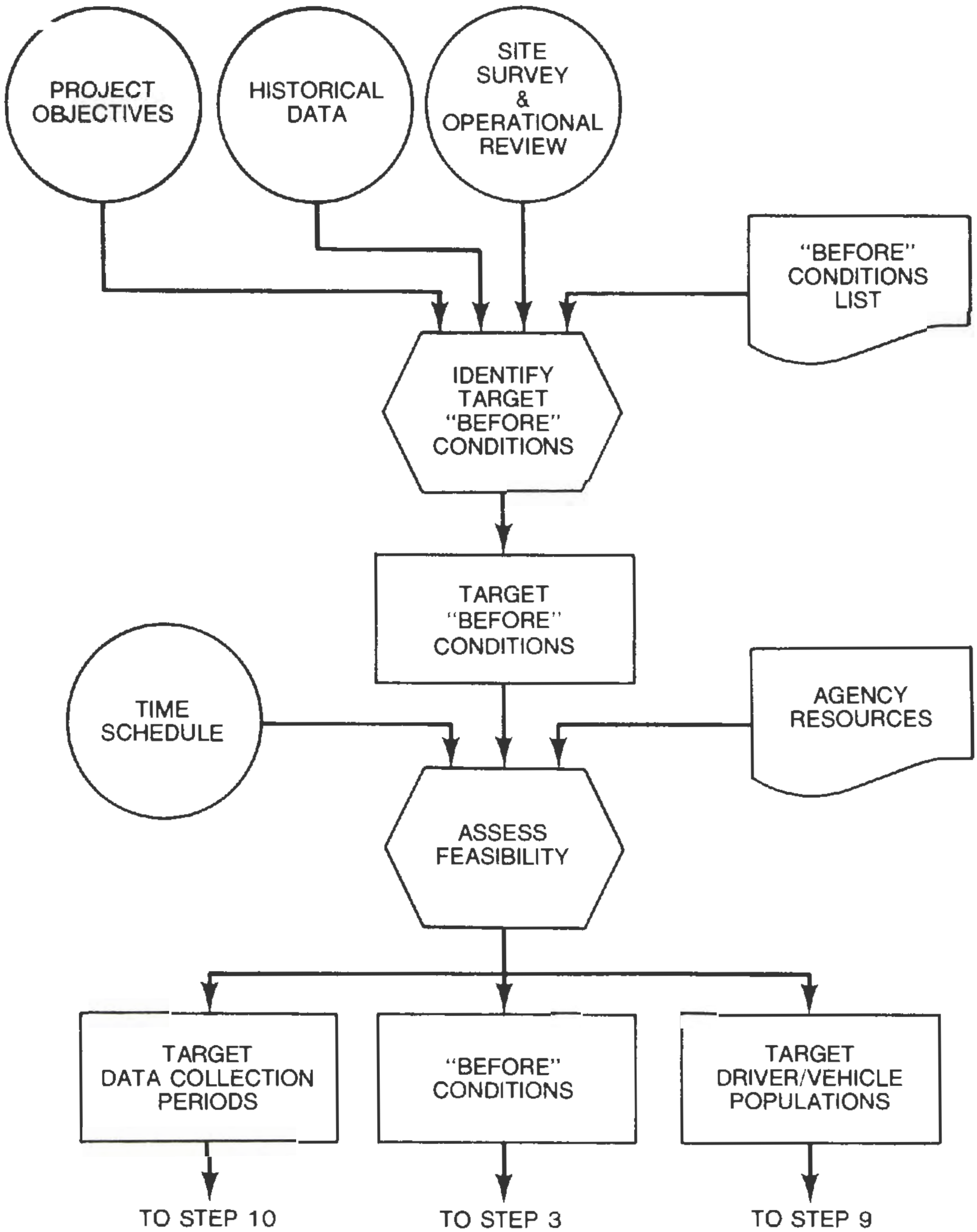


Figure 2-3.
Flow, Step 2 - Specify "Before" Conditions

"BEFORE" CONDITIONS —STEP 2		
PROJECT:	*I 293 Lane Drop at Sylvan Rd. Exit*	
PROJECT NO.:	*AI-1*	
EVALUATOR/DATE:	*HL 3/9/80*	
CONDITIONS:	WEATHER:	*Clear*
	ILLUMINATION:	*Day time*
	PAVEMENT:	*Dry*
	VOLUME:	*20 K*
	SEASON:	*Spring, Summer, Fall*
	OTHERS:	*Trees–Full green foliage* *New Striping* *Beach traffic at peak*
TARGET POPULATIONS:		
	DRIVER MIX:	*Tourist population*
	VEHICLE MIX:	*80% Passenger Cars, 10% R-V's, 10% Trucks*
TARGET DATA COLLECTION PERIODS:		
	TIME(S) OF DAY:	*P.M. peak*
	DAY(S) OF WEEK:	*Weekends, holidays*
	SEASON:	*Summer tourist season*
COMMENTS	*Tourists are getting pulled off at lane drop*	

Exhibit 2-2.
Sample "Before" Conditions Listing Form

STEP 3. Select Measures of Effectiveness (MOE's)

Measures of Effectiveness (MOE's) used to diagnose problems and evaluate improvements should be directly related to project objectives. MOE's may be classified by accident, traffic performance, and system performance categories. Accident MOE's are expressed in terms of rate and calculated frequency. Traffic performance MOE's consist of measurable movements of vehicles' speed and path, and also include surrogate measures such as erratic maneuvers and conflicts. System performance MOE's are basically measure traffic flow parameters such as service volume (to measure capacity) and delay.

Just as there usually are several objectives associated with a project, there will generally be several MOE's that can be used to determine the project's effectiveness. Given that all projects gather data to diagnose problems, most diagnostic measures also serve as effectiveness measures, particularly when objectives relate to traffic and system performance improvements. Even when a project's primary objective is accident reduction, traffic performance and system performance MOE measures should also be used. These measures will aid in generating improvements and may eliminate methodological (e.g., poor accident records system) or logistical (e.g., long time frame) problems associated with the sole use of accident MOE's.

MOE's must be: (1) Directly related to the objectives, i.e., valid; (2) stable and repeatable, i.e., reliable; (3) amenable to data collection--given a particular equipment/manpower situation, i.e., feasible; and (4) of value in diagnosing a problem, i.e., meaningful. Validity, reliability, feasibility and meaningfulness are the primary MOE selection criteria.

Validity relates to what is measured. It is established on the basis of a measure's ability to predict future performance. The best way to insure that an MOE is valid is to select one that has already been validated in previous similar projects. Reliability relates to how data is collected. It is established on the basis of consistent, repeatable measurements. The best way to insure reliability is to use calibrated, stable equipment and trained personnel.

Feasibility relates to whether and at what cost data can be collected. It is established on the basis of the kind and availability of equipment and personnel and on the characteristics of the site. The best way to insure feasibility is to only choose MOE's which match the resources of the performing agency and are suitable for data collection at the site. Meaningfulness relates to what if a change is obtained and how it is interpreted. It is established on the basis of practical usefulness and sensitivity in diagnosing a problem. The best way to insure that an MOE is meaningful is by engineering judgment. For example, if all vehicles were applying their brakes, then brake light applications might not be a meaningful MOE.

In order to select MOE's, the evaluator should consider each objective and category to identify the type of MOE that could be used. Table 2-4 presents a number of validated measures for typical highway situations. In addition, MOE's have been used to evaluate the traffic control devices shown in Table 2-5. While specific improvements will not have been generated at this juncture in the project, personnel generally have a feel for the types of traffic control devices that will be used. Thus, using the site's characteristics, potential solutions, and engineering judgment will yield a range of candidate MOE's.

After MOE's have been selected, exposure units, when applicable, should be specified. Exposure relates to the quantity of vehicles, vehicle miles or other volume and/or time related factors which measure the degree of vehicle or driver exposure to a particular MOE. Most system and traffic performance MOE's are "categorical." Their exposure units are rate-related. Other MOE's, such as speed, delay, lateral placement, passing time and time through an intersection are referred to as "continuous," and exposure units are not applicable for these measures. For those projects where accident MOE's are used, their exposure units may be expressed as rate (usually per hundred million vehicle miles) or in terms of frequency, using accidents per year as the exposure unit. When rate-related MOE's are used, exposure is expressed per unit volume (usually per 1,000 vehicles) for spot or short segment improvements, or in terms of vehicle-miles for extended segments. It is, therefore, always necessary to obtain volume counts for rate-related MOE's.

TABLE 2-4. MOE'S FOR TYPICAL HIGHWAY SITUATIONS

SITUATION	MOE's	SITUATION	MOE's
ALIGNMENT, HORIZONTAL CURVE	Spot Speed; Upstream; Entry; Apex; Exit; Downstream. Lateral Placement. Encroachments; Shoulder; Centerline. Brake Applications.	MERGE	Merge speed profile. Conflicts with through stream. Distribution of merges. Delay. Brake Applications.
ALIGNMENT, VERTICAL CURVE	Spot Speed; Upstream; Entry; Sag (or); Crest; Downstream. Brake Applications. Time Headway (Downgrade).	NARROW BRIDGES	Speed. Lateral Placement. Centerline Encroachment. Conflicts.
CONSTRUCTION AND MAINTENANCE ZONES	Brake Applications. Conflicts. Delay. Encroachments. Lateral Placement. Last Minute Lane Change. Speed.	OBSTACLES	Speed. Lateral Placement.
CHANGE IN CROSS SECTION- Lane, Shoulder Width Reduction	Spot Speed. Brake Applications. Lateral Placement.	PASSING ZONES	Passing frequency. Passing and return type. Number of abortive passes. Conflicts with oncoming or overtaken vehicles.
INTERCHANGES, DIVERGE AREAS	Distribution of points of entry into inside lane; decel. lane. Speed; mainstream (reduction); at gore area; on ramp. Decel. lane speed profile. Erratic Movements.	PEDESTRIAN CROSSING: SCHOOL CROSSING	Compliance. Conflicts. Speed. Delay.
INTERCHANGES, LEFT EXITS, TANGENTIAL OFF RAMPS	Erratic Movements. Conflicts. Speed. Lateral Placement	RAILROAD CROSSING	Head Turning Movements. Speed. Speed Profile.
INTERSECTIONS, SIGNALIZED	Conflicts. Delay. Travel Time. Time Through Intersection. Speed. Lateral Placement. Brake Applications. Stop Line Encroachments.	STOP APPROACH	Speed Profile. Lateral Placement. Brake Applications. Encroachments on Cross Roadway. Erratic Deceleration.
LANE DROPS	Spot Speed; Upstream; Vicinity of sign; Beginning taper; End taper. Distribution of lane changes. Merging conflicts. Encroachments. Lateral Placement through transition area.	TURN	Location of lane changes to enter decel. lane. Spot speeds; Upstream; Entry. Point of entry into decel. lane. Erratic maneuvers. Conflicts; opposing; through vehicles. Time through intersection.
		TOLL PLAZAS	Speed. Lateral Placement. Conflicts.
		WEAVING SECTIONS	Speed. Speed Change. Brake Light Applications. Conflicts. Lateral Placement.

TABLE 2-5. MOE'S FOR TYPICAL TRAFFIC CONTROL DEVICES

DEVICE	MOE
MARKINGS	. Lateral Placement . Encroachments . Compliance
GUIDE SIGNS	. High Risk Gore Weaves (Erratic Maneuvers) . Gore Weaves (Erratic Maneuvers) . Driving Slowly . Late Lane Changes . Brake Light Applications . Energy Efficiency
WARNING SIGNS	. Speed (profile) (Spot) . Lateral Placement . Brake Light Indications . Stop Line Conflicts . Compliance
SIGNALS	. Conflicts . Speed . Compliance . Energy Efficiency

The Step 3 flow is shown in Figure 2-4. MOE's should be selected for each objective and entered on the MOE Listing form shown in Exhibit 2-3. There should be a separate listing for each objective.

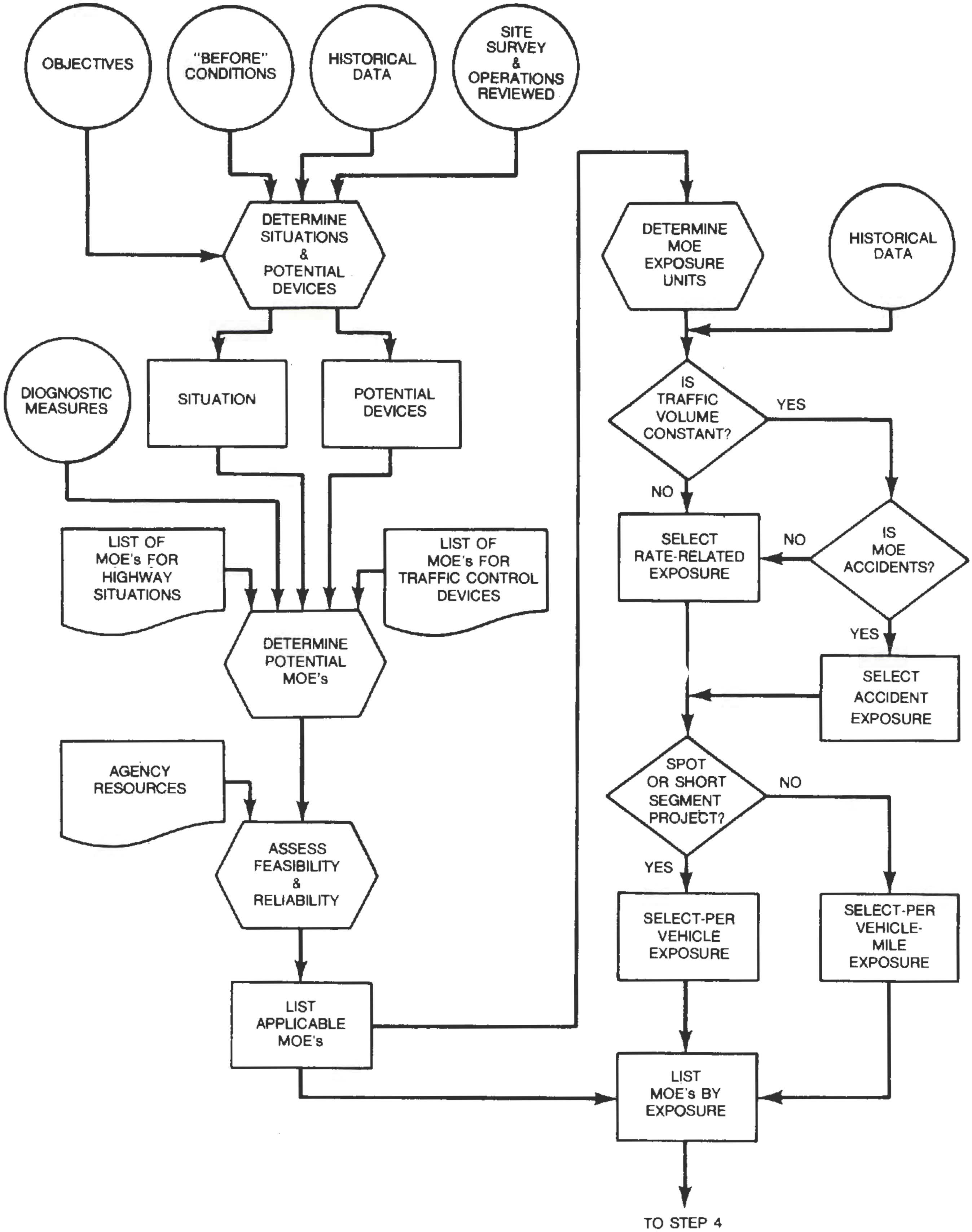


Figure 2-4.
Flow, Step 3 - Select Measures of Effectiveness

MOE LISTING—STEP 3		
PROJECT: *I 9 & I 309 Split*		
PROJECT NO.: *A 47*		
EVALUATOR/DATE: *HL 3/7/80*		
OBJECTIVE: *Traffic Performance – To reduce erratic maneuvers and speed at gore.*		
SITUATION: *Diverge Area – 4:2:2 split-off route to left of through route.*		
POTENTIAL TRAFFIC CONTROL DEVICE(S): *Major guide sign change expected.*		
MOE's	MOE TYPE	EXPOSURE UNIT
Erratic Maneuvers – (High risk gore weave)	*Categorical*	*Per 1000 vehicles*
Speed (at gore)	*Continuous*	*N/A*

Exhibit 2-3.
Sample MOE Listing Form

STEP 4. Select Measurement Technique and Operationally Define MOE's

The main emphasis in Step 3 was validity and meaningfulness. In this step, feasibility and reliability are addressed. Table 2-6 shows applicable data collection techniques for each MOE in terms of actuarial, observational, and interactive methods.

Actuarial - An actuarial data collection method uses historical records to quantify MOE's in terms of their past frequency or rate of occurrence.

Observational - This method gathers data by observing (usually unobtrusively) driver and/or vehicle behavior as it is occurring. The personnel and/or equipment used to gather data should not interact with what is being observed.

Interactive - In this method, the individual and/or equipment interacts directly with the drivers whose opinions, understanding or knowledge is being solicited, or whose behavior is being observed.

Using Table 2-6, an assessment should be made of the present or future availability of equipment and staff (see Report FHWA-TO-80-2). This, coupled with an assessment of the site's characteristics, yields an indication of the feasibility of using a particular MOE and measurement technique. For example, lack of equipment and personnel to reliably measure lateral placement would preclude its use as an MOE. Similarly, if there is no suitable vantage point to mount a camera, time-lapse techniques to record erratic maneuvers would not be feasible.

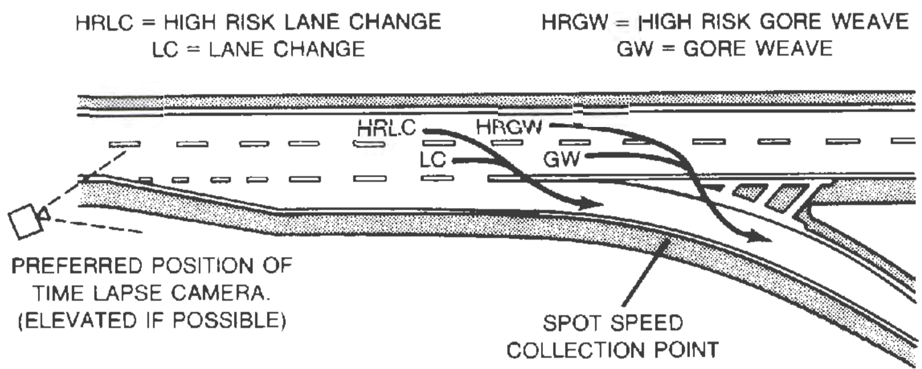
Each MOE should then be operationally defined. An operational definition states the operations and/or procedures employed in distinguishing the MOE from other measures. It defines the MOE in precise, observable terms and reduces the chance for error in data collection, reduction and analysis. For example, an MOE such as "driving slowly" might imply different things to different observers. An operational definition of driving slowly may be as follows: "a vehicle speed equal to or less than one standard deviation (e.g., 5 mph) below the mean speed (e.g., 57 mph) measured 800 feet in advance of the physical gore." An operational definition such as this would eliminate any ambiguity. It will result in the same data being taken in the "Before" and "After" phase, and will also dictate the location of the observer and equipment and the required accuracy of the data collection method.

TABLE 2-6. APPLICABLE DATA COLLECTION TECHNIQUES

MOE	DATA COLLECTION TECHNIQUE	METHOD
ACCIDENTS	Accident Records	Actuarial
BRAKE APPLICATIONS	Manual Recording Time-Lapse Film Traffic Analyzer Traffic Counter Video Recorder Wear Patterns Interview Questionnaire	Observational Observational Observational Observational Observational Observational Interactive Interactive
COMPLIANCE	Police Records Manual Recording Time-Lapse Film Traffic Counter Interview Questionnaire	Actuarial Observational Observational Observational Interactive Interactive
CONFLICTS	Manual Recording Time-Lapse Film Traffic Counter Video Recorder	Observational Observational Observational Observational
DELAY	Complaints Aerial Photography Automatic Detectors Input-Output Studies Manual Recording Moving Vehicles Radar Speed Meters Time-Lapse Film Traffic Analyzer Traffic Counter Video Recorder Interview Questionnaire	Actuarial Observational Observational Observational Observational Observational Observational Observational Observational Observational Observational Interactive Interactive
DRIVING SLOWLY	Aerial Photography Input-Output Studies Manual Recording Moving Vehicles Radar Speed Meters Time-Lapse Film Traffic Analyzer Video Recorder	Observational Observational Observational Observational Observational Observational Observational Observational
ENCROACHMENTS	Accident Records Traffic Counts Aerial Photography Automatic Detectors Manual Recording Moving Vehicles Time-Lapse Film Traffic Analyzer Video Recorder Wear Patterns Interviews Questionnaire	Actuarial Actuarial Observational Observational Observational Observational Observational Observational Observational Observational Interactive Interactive

MOE	DATA COLLECTION TECHNIQUE	METHOD
ERRATIC MANEUVERS, LAST MINUTE LANE CHANGES	Accident Records Aerial Photography Automatic Detectors Manual Recording Moving Vehicles Time-Lapse Film Traffic Analyzer Video Recording Interviews Questionnaire	Actuarial Observational Observational Observational Observational Observational Observational Observational Interactive Interactive
HEAD TURNING MOVEMENTS	Manual Recording Time-Lapse Film Video Recording	Observational Observational Observational
LATERAL PLACEMENT, MERGES	Aerial Photography Automatic Detectors Manual Recording Moving Vehicles Time-Lapse Film Traffic Analyzer Video Recording	Observational Observational Observational Observational Observational Observational Observational
PASSING TYPE, PASSING TIME, PASSING DISTANCE, PASSING FREQUENCY, PASSING ABORTIVE	Aerial Photography Manual Recording Moving Vehicles Time-Lapse Film Traffic Analyzer Video Recording	Observational Observational Observational Observational Observational Observational
SPEED, SPEED CHANGES	Accident Records Aerial Photography Input-Output Studies Manual Recording Moving Vehicles Radar Speed Meters Time-Lapse Film Traffic Analyzer Video Recorder	Actuarial Observational Observational Observational Observational Observational Observational Observational Observational
TIME HEADWAY	Aerial Photography Automatic Detectors Manual Recording Moving Vehicles Time-Lapse Film Traffic Analyzer Video Recorder	Observational Observational Observational Observational Observational Observational Observational
TIME THROUGH INTERSECTION, TRAVEL TIME	Aerial Photography Automatic Detectors Input-Output Studies Manual Recording Moving Vehicles Radar Speed Meters Time-Lapse Film Traffic Analyzer Video Recording	Observational Observational Observational Observational Observational Observational Observational Observational Observational
VOLUME	Traffic Counts Aerial Photography Automatic Detectors Input-Output Studies Manual Recording Moving Vehicles Time-Lapse Film Traffic Analyzer Video Recording	Actuarial Observational Observational Observational Observational Observational Observational Observational Observational

Table 2-7 presents functional definitions for a range of MOE's. These should be translated into operational definitions by synthesizing site characteristics and data collection techniques and procedures with the MOE's. It should be recognized that an operational definition is always site specific and tailored to the project location. Figure 2-5, taken from Hanscom and Berger (1976), shows examples of these interrelationships.



MOE.	HIGH RISK GORE WEAVE	GORE WEAVE	LATE LANE CHANGE	DRIVE SLOWLY
OPERATIONAL DEFINITION:	A VEHICLE MOVEMENT INTO DECELERATION LANE ACROSS PAINTED OR PHYSICAL GORE, IN ADDITION TO CROSSING AT LEAST ONE THROUGH TRAFFIC LANE.	A VEHICLE MOVEMENT INTO DECELERATION LANE ACROSS PAINTED OR PHYSICAL GORE.	A VEHICLE MOVEMENT INTO DECELERATION LANE ACROSS PAINTED GORE EXTENSION LINE.	A VEHICLE SPEED < ONE STANDARD DEVIATION BELOW MEAN. 800 FEET IN ADVANCE OF PHYSICAL GORE POINT.
COLLECTION METHOD:	TIME LAPSE PHOTOGRAPHY	MANUAL CODING* OR TIME LAPSE *MANUAL CODING IS PREFERABLE IF TOTAL WEAVE AREA IS 1000 FEET OR LONGER.	TIME LAPSE PHOTOGRAPHY	MANUAL TIMING VIA ELECTRONIC STOPWATCH
COLLECTION PROCEDURE:	MEASURE OR COUNT ALL OCCURRENCES CONTINUOUSLY			SLOW AND MEAN SPEED DURING ALTERNATE PERIODS.

Figure 2-5.
Interrelationships of MOE's and Methods

TABLE 2-7. DEFINITIONS FOR SELECTED MOE's

MOE	Definition
CONFLICTS, TRAFFIC (overall definition)	A traffic event involving two or more road users, in which one road user performs some atypical or unusual action, such as a change in direction or speed, that places another user in jeopardy of a collision unless an evasive maneuver is undertaken.
(Nine Basic Intersection Conflicts)	(See FHWA-TO-80-2 for applicable Figures)
1. Left-turn, same direction	An instigating vehicle slows to make a left turn, placing a following, conflicted vehicle in jeopardy of a rear end collision. The conflicted vehicle brakes or swerves, then continues through the intersection.
2. Right-turn, same direction	Same as (1) above except instigating vehicle slows to make a right turn.
3. Slow-vehicle, same direction	Same as (1) above except instigating vehicle slows while approaching or passing through an intersection.
4. Opposing left-turn	An oncoming vehicle makes a left turn, placing the conflicted vehicle in jeopardy of a head-on or broadside collision. The conflicted vehicle brakes or swerves, then continues through the intersection. The conflicted vehicle is assumed to have the right-of-way in this and subsequent conflict categories.
5. Right-turn, cross-traffic-from-right	An instigating vehicle approaches from the right to make a left turn, placing the conflicted vehicle in jeopardy of a broadside or rear end collision. The conflicted vehicle brakes or swerves, then continues through the intersection.

TABLE 2-7 (continued)

MOE	Definition
6. Left-turn, cross-traffic-from-right	An instigating vehicle approaching from the right makes a left turn, placing the conflicted vehicle in jeopardy of a broadside collision. The conflicted vehicle brakes or swerves, then continues through the intersection.
7. Thru, cross-traffic-from-right	Same as (6) above except instigating vehicle approaching from right comes in front of the conflicted vehicle.
8. Left-turn, cross-traffic-from-left	Same as (5) above except instigating vehicle approaching from the left makes a left turn.
9. Thru, cross-traffic-from left	Same as (6) above except instigating vehicle crosses in front of a conflicting vehicle.
DRIVING SLOWLY	A vehicle speed greater than X standard deviation(s) below the mean speed, measured Y feet in advance of the physical gore.
DELAY	There are a number of kinds of delay - The two most prevalent are Approach and Stopped time. Each is measured by the "Point Sample, Stopped Delay Method" (See Reilly, et al., 1976).
Approach Delay	The total amount of time, in vehicle seconds, lost by vehicles due to traffic conditions on the approach to a signalized intersection. For an individual vehicle, approach delay is considered to be the amount of time used to pass through an approach minus the amount of time used by an unimpeded vehicle to pass through the approach.
Stopped Time Delay	The time, in vehicle seconds, during which a vehicle is stopped with locked wheels on the intersection approach.

TABLE 2-7 (continued)

MOE	Definition
ENCROACHMENTS	Deviation from a prescribed path in a lane, intersection, or interchange as described by the existing pavement markings. An encroachment occurs when a wheel or wheels of an encroaching vehicle touches or goes across a lane line, centerline, edgeline, or other feature.
ERRATIC MANEUVERS	Deviations from an idealized track or trace through an interchange given a particular destination. The paths should be defined for both exiting and through traffic, and will differ for different types of interchange geometry. At locations up-stream of interchanges, preparatory and through maneuvers must be defined (see FHWA-TO-80-2 for applicable Figures).
Late Lane Change	A vehicle movement into the deceleration lane across the painted gore extension lane.
Gore Weave	A vehicle movement into the deceleration lane across the painted or physical gore.
High Risk Late Lane Change or Gore Weave	Same as above with the addition of crossing at least one through traffic lane.
PREPARATORY MANEUVER	Preparation for exiting, e.g., moving into the right lane for a right exit, etc.
THROUGH MANEUVER	Moving out of a lane, e.g., moving into the left lane at a right exit, etc.

A flow chart for Step 4 is presented in Figure 2-6. The MOE Definition form is shown in Exhibit 2-4.

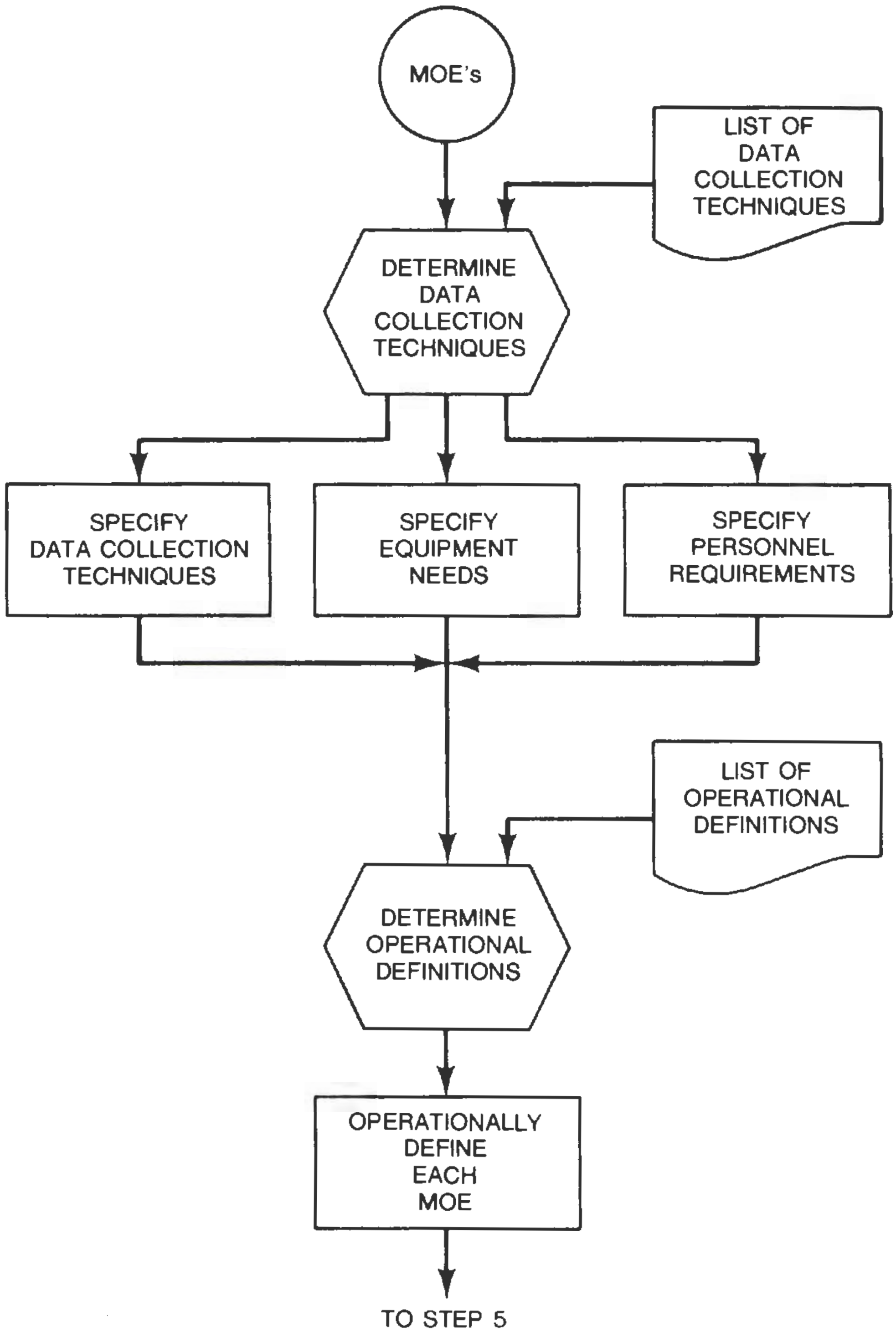


Figure 2-6.
Flow, Step 4 - Select Measurement Technique and Operationally Define MOE's

MOE DEFINITION—STEP 4

PROJECT: *Smith Street Interchange–Westbound*

PROJECT NO.: *Z 97*

EVALUATOR/DATE: *HL 2/18/80*

MOE: (1) *Erratic Maneuvers–High risk gore weave; gore weave; late large change*

(2) *Driving Slowly*

DATA COLLECTION TECHNIQUE(S):

(1) *Time lapse*

(2) *Manual timing–electronic stop watch*

(3) *Possible manual coding as back-up technique*

EQUIPMENT/PERSONNEL REQUIREMENTS:

(1) *Minolta D-210 Camera (Super-8)*
(2) *Camera Mount*
Super-8 Film

(1) (2) *2 Data Collection people–*
1 electronic stop watch

(3) *2 Data Collection people*
Clipboard
Pencils & Forms

Exhibit 2-4.
Sample MOE Definition Form

MOE DEFINITION

OPERATIONAL DEFINITION:

Gore Weave–A movement in the decel lane across the painted or physical gore from the adjacent lane.

High Risk Gore Weave–S/A Gore Weave except from median lane.

Lane Change–A vehicle movement in the decel lane across the painted gore extension lane.

High Risk Lane Change–Same as Lane Change except from median lane.

Drivers Slowly–A vehicle travels at 40 MPH or less 800 feet in front of physical gore.

SKETCH:

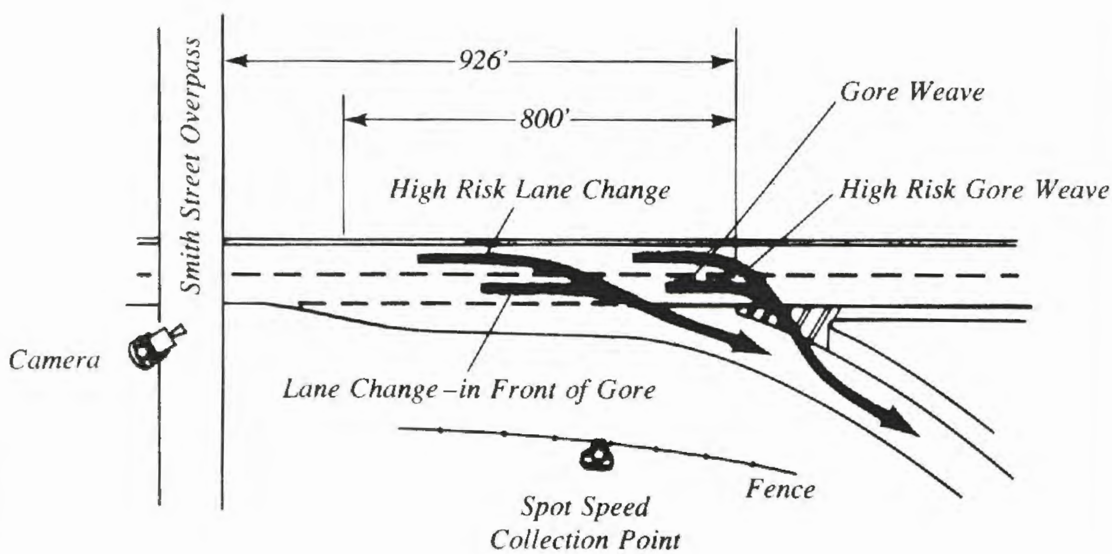


Exhibit 2-4.
(Continued)

STEP 5. Specify Implementation and Acclimation Period

The Implementation and Acclimation period encompasses the time from the end of "Before" data collection to the beginning of "After" data collection. The implementation portion represents the time it takes to develop, design, fabricate, schedule, and apply the improvement. Its length is dependent on a number of factors relating to the improvement developmental process and the nature of the improvement. Availability of resources, personnel, equipment, etc., as well as agency policy, also affects the time frame of the implementation portion. The acclimation portion represents the time it takes repeat drivers to become accustomed to the improvement. If the time between implementation and "After" data collection were not allowed to pass, differences in the MOE's could be attributable to the "novelty effect," i.e., a response to a change per se rather than to the improvement.

There is no acclimation portion for accident MOE's, since "After" data are collected immediately after installation of the improvement. Thus, the length of the Implementation and Acclimation period for this class of MOE's is solely the implementation portion. However, since "After" data are required to be accumulated for a 3-year period following implementation, it should be recognized that the overall time frame for the project evaluation using accident MOE's will be over 3 years. This is based on the assumption that 3 years of suitable "Before" accident data exist.

Specifying the length of the implementation and acclimation period often involves several tradeoffs. While the implementation portion should be as brief as possible, external factors and agency policy often dictate its length. On the other hand, the acclimation portion must be long enough to eliminate the novelty effect. However, the total Implementation and Acclimation period should be short enough to minimize changes over time.

Characteristics that may change over time include traffic volume, traffic stream mix, travel patterns, road surface characteristics, and enforcement procedures. One way to control for these changes is to specify as brief an intervening period as possible. Other ways to control for changes include: Anticipating changes and avoid starting the evaluation, e.g., when construction is scheduled, or when a new traffic generator is anticipated; shortening the acclimation portion if changes are occuring; and eliminating changes by coordination with other jurisdictions or agencies, e.g., coordinating with the police to maintain consistent enforcement. If changes over time are likely, and if it is not possible to specify a short acclimation portion, then consideration might be given to using the Control Site design (see Step 6).

Other factors that enter into specifying the length of the Implementation and Acclimation period involve comparability. The conditions under which the "After" data are collected must be comparable to the "Before" conditions specified in Step 2. Consideration must also be given to controlling for seasonal variations in such things as climate, hours of daylight, landscaping, foliage and the ratio of tourists to locals. The specification of equipment measurement accuracy and the development of good operational definitions will also help assure comparability. In addition, the equipment and personnel used in the "Before" and "After" phases should ideally be the same, but at worst, must be comparable. Thus, their availability often dictates the length of the period.

For traffic and system performance MOE's, the acclimation portion should be a minimum of 30 days following improvement implementation. For many projects, a 1-year Implementation and Acclimation period has been found to be optimum particularly in situations where there are navigational changes (e.g., new guide signs) since it is long enough to eliminate the novelty affect, short enough so that there is not likely to be any major changes over time, and just the right period to assure comparability of seasonal conditions.

Specifiying the Implementation and Acclimation period thus requires a number of decisions. When accident MOE's are used, no acclimation portion is required, but 3 years of "After" data must be accumulated (data may be assessed for trends after the first and second year). This is required to counter the effects of "regression toward the mean" (spontaneous return to an average accident frequency from an unusual peak), and to accumulate a large enough data base to permit application of statistical tests. The "After" data collection period should be comparable to the "Before" period (month, day). For other MOE's, a 30-day minimum acclimation portion following improvement implementation is required.

In all instances, adjustments should be made to insure comparability, particularly if there are seasonal variations in climate, hours of daylight, landscape and foliage, traffic system make-up, etc. If navigational issues are relevant, a 1-year total period should be specified. A determination should also be made of whether there will be any possible changes in the site during the acclimation portion, and what controls could be applied. Table 2-8 provides a listing of changes and controls that can be used. The Step 5 flow is shown in Figure 2-7. Exhibit 2-5 shows a sample form.

TABLE 2-8. CHANGES OVER TIME AND WAYS TO CONTROL FOR THE CHANGES

CHANGE	CONTROL
Volume (increase or decrease)	Shorten acclimation portion.
Climate (Seasonal variations)	Adjust Implementation and Acclimation period to insure similar conditions.
Enforcement	Coordinate with police to assure comparability.
Geometric changes (minor construction)	Do not start construction until evaluation is complete.
Maintenance	Adjust acclimation portion to be finished before maintenance or delay maintenance.
New traffic Generators	Shorten acclimation portion or coordinate to delay new generators

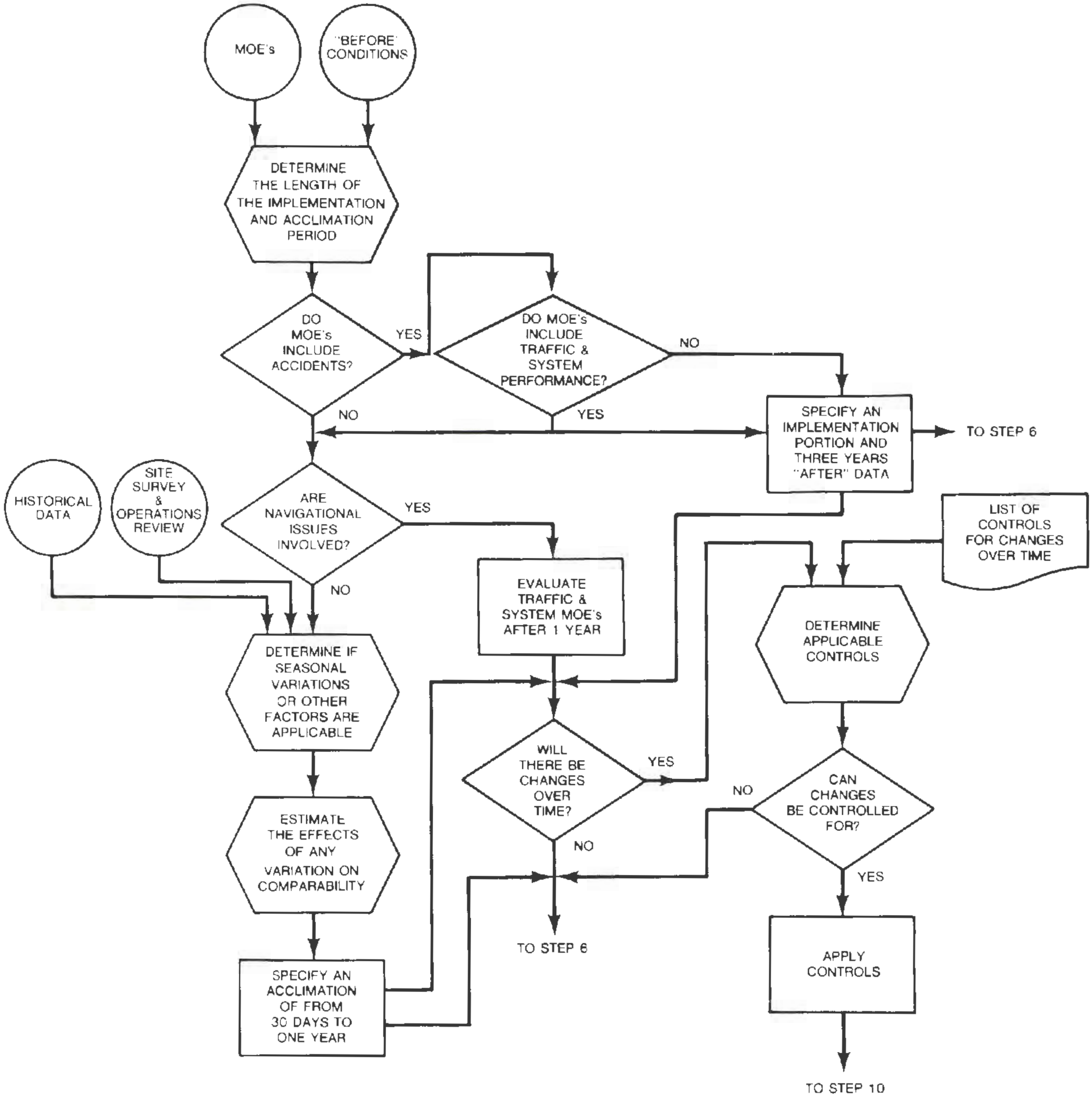


Figure 2-7.
Flow, Step 5 - Specify Implementation and Acclimation Period

IMPLEMENTATION AND ACCLIMATION PERIOD—STEP 5

PROJECT: *Interchange of 7th Street Bypass*

PROJECT NO.: *1234 B*

EVALUATOR/DATE: *HL 6/1/79*

MOE CATEGORIES: ☐ ACCIDENT REDUCTION

☑ TRAFFIC/SYSTEM PERFORMANCE

ESTIMATED START OF IMPLEMENTATION PORTION: *July 17, 1979*

ESTIMATED START OF ACCLIMATION PORTION: *October 12, 1979*

POSSIBLE CHANGES OVER TIME	CONTROLS
– *Enforcement*	*Coordinate with State H.P.*
– *Main Street Mall Development*	*Coordinate with Planning*
– *Volume change next summer – increase due to summer season*	*Shorten acclimation to < 1 year*

TOTAL PERIOD(S):

☐ ACCIDENT MOE's: ________ PLUS 3 YEARS "AFTER" DATA ACCUMULATION

☑ 1 YEAR

☑ OTHER: *8-9 months*

ESTIMATED END OF PERIOD:

July 1980 – Possible data collection in April 1980 if needed.

Exhibit 2-5.
Sample Implementation and Acclimation Period Form

STEP 6. SELECT EVALUATION DESIGN

The evaluation design (experimental design) is the plan used to determine the effectiveness of the improvement. It serves as the analytical framework for: The selection of samples; the kind, order and procedure for the administration of the improvement; and the recording and statistical analysis of the data.

Two designs are recommended. The "Before-After" and the "Before-After with a Control Site" (Control Site design). The design used depends on a number of factors related to the advantages and disadvantages of each.

Before-After - The Before-After design is shown schematically in Figure 2-8. When this design is used, MOE data are collected at the project site prior to the improvement, and, after a suitable acclimation period, following the application of the improvement. Differences in the MOE's before and after the improvement (which represent the period averages) are evaluated for statistical significance. This design is based on the following assumptions: (1) The MOE levels before the application of the improvement would have remained at the same level following the implementation and acclimation period if no improvement were applied; (2) the only thing changed at the site is the improvement; and (3) any differences in the MOE's are attributable to the improvement (comparisons in the "After" period are actually between the "expected after" and "actual after").

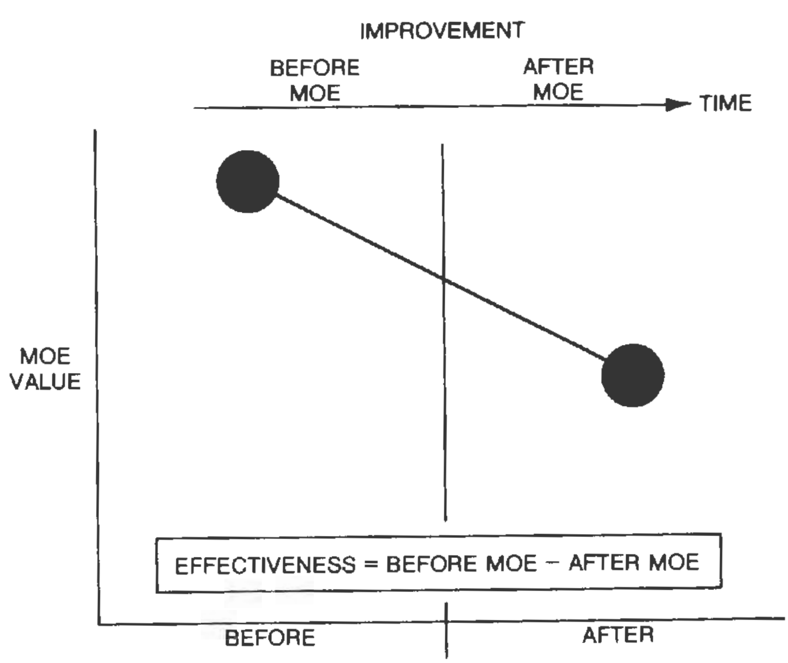


Figure 2-8.
Before - After Design

Before-After with a Control Site - Figure 2-9 presents a schematic of the Control Site design. This design is similar to the Before-After design with the exception of the inclusion of an additional site or sites, selected to serve as control sites. MOE data are collected at the project and control sites at the same time during the "Before" phase. The improvement is then only applied to the project site. Following identical implementation and acclimation periods, "After" data are collected at both the project and control sites. Differences in the "After" phase MOE's between the project and control site are evaluated for statistical significance. This design is based on the following assumptions: (1) Project and control sites are comparable; (2) MOE levels at project and control sites are similar; (3) while the only changes at either site should be the application of the improvement at the project site, if there are any changes over time, they will affect both the project and the control sites; and (4) differences in the MOE's that can be attributed to the improvement can be determined by factoring out any changes over time found to occur at the control sites.

SITE	BEFORE	IMPROVEMENT	AFTER
CONTROL	MOE	NONE	MOE CONTROL
PROJECT	MOE	IMPROVEMENT	MOE PROJECT

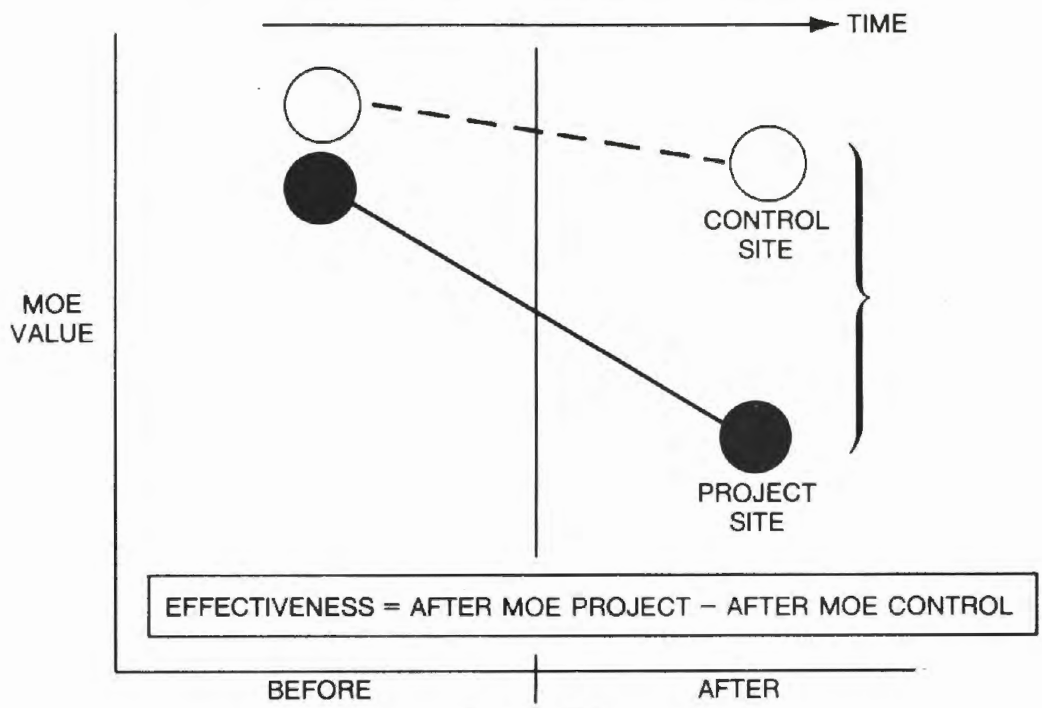


Figure 2-9.
Before - After with a Control Site Design

The Before-After design is suitable for evaluating most projects. This design may be used when: (1) A jurisdiction lacks the personnel, equipment and resources to use a Control Site design; (2) the time period between the "Before" and "After" data collection phase is relatively short, such as when traffic and system performance MOE's are being evaluated; (3) a rate-related MOE is used, since it washes out the effects of changes in traffic volume; and (4) there is no suitable control site.

When a Control Site design is deemed advisable, generally for safety projects where a long time period must intervene between the "Before" and "After" phases, and when control site(s) are available, the control site(s) must be comparable. If a control site is not comparable, then the conclusions are more likely to be erroneous than if no control site were used at all. Table 2-9 lists criteria to be used in selecting suitable control sites. When long time periods intervene, either linear regression (when there are trends apparent) or point value averages (when no trend is apparent) must be used to compute and compare "After" MOE values (see Appendix B).

TABLE 2-9. CONTROL SITE SELECTION CRITERIA

1.	Select geometrically similar site(s).
2.	Select sites with similar traffic control devices.
3.	Check whether MOE's at control site(s) are within ±10 percent* of the project site.
4.	Check whether other key variables such as traffic at control site(s) are within ±10 percent* of the project site.
5.	Check that there are no plans to modify the control site(s) during the project.
6.	Check that lighting, vehicle mix and land use are similar.
*	The ±10 percent criteria is a guide

A flow for Step 6 is given in Figure 2-10. Exhibit 2-6 presents a sample Evaluation Design listing form.

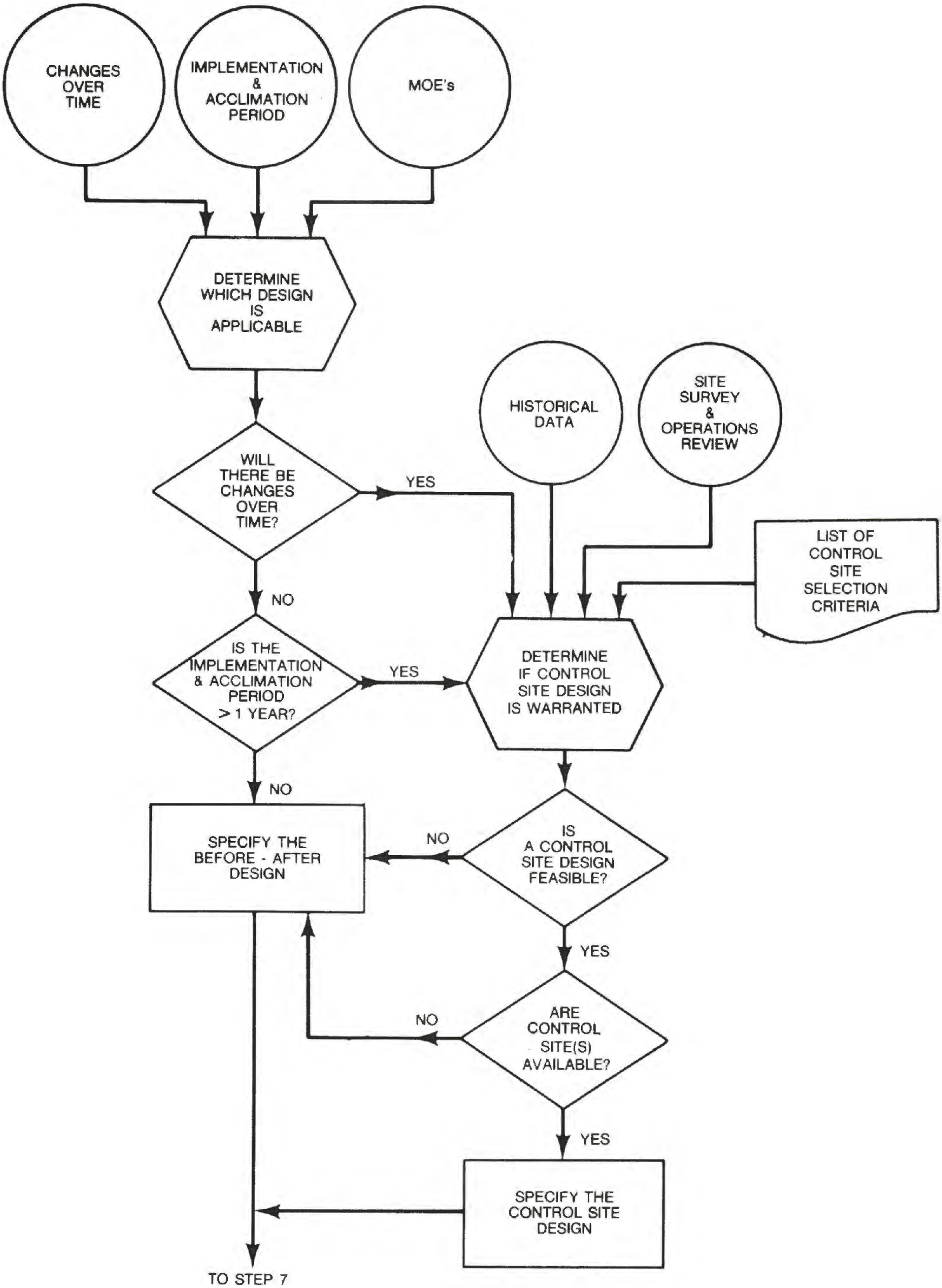


Figure 2-10.
Flow, Step 6 - Select Evaluation Design

EVALUATION DESIGN—STEP 6

PROJECT: *Intersection of Washington & Jefferson Streets S.E.*

PROJECT NO.: *AA-17*

EVALUATOR/DATE: *HL 3/17/80*

DESIGN:

☐ BEFORE-AFTER

☑ BEFORE-AFTER WITH A CONTROL SITE

FOR BEFORE-AFTER WITH A CONTROL SITE DESIGN, IDENTIFY CONTROL SITE(S):

Washington Ave. & Hamilton Ave. S.W.

COMMENTS:

Control site is 17 blocks from project site. It has the same geometry, land use, etc.

During the acclimation, the team must monitor volume since traffic may be shifting when Smithfield Mall is built (Spring 1982).

Exhibit 2-6.
Sample Evaluation Design Form

STEP 7. Specify Statistical Tests

Effectiveness is inferred on the basis of the validity of the evaluation, and the attainment of practical and statistically significant results. The first six steps serve to structure a valid evaluation design. This and the next step deals with how to determine whether results are statistically significant. Before specifying a statistical test, it would be useful for the reader to refer to Appendix A for a brief discussion of relevant descriptive and inferential statistical factors.

Statistical methodology assumes an important role throughout the evaluation, since statistics are used to plan data collection, to tabulate MOE's and to analyze results. Descriptive statistics organize, summarize and describe the data. Inferential statistics test for significance, thereby enabling evaluators to go beyond what is measured to estimate how close the data are to the "real world" and to establish how meaningful the results are for forecasting and decisionmaking.

The determination of whether an improvement is effective ultimately rests on the results of an inferential statistical analysis of differences in the MOE before and after an improvement is applied. At issue is whether differences are a result of the improvement or could have occurred by chance. Significance is established by testing whether the difference could have occurred significantly often by chance. If this can be shown to be false, then the result is attributed to the improvement.

A number of statistical tests have been developed to test for significance. These tests are designed for use with different evaluation designs, data types, and distributions. Some are very restricted in their application, while others are applicable to a broad range of conditions. The selection of an appropriate test rests, to a large extent, on how closely the data fits the assumptions of the test. Table 2-10 shows data types for various MOE's.

TABLE 2-10. DATA TYPES

Continuous data, such as speeds and delay from which means, percentiles, or standard deviations can be obtained.
Categorical or count data, such as erratic maneuvers and traffic conflicts from which rates or proportions are derived.
Rare event data, such as accidents, from which rates and yearly magnitudes are determined.

Five tests of significance are recommended for use in the evaluation. This is not meant to imply that there are no other tests which can be used. Those familiar with statistical methodology may feel that another test is more appropriate. However, when a test not recommended by this report is used, its use should be justified and its assumptions spelled out. The five recommended tests are:

- Poisson
- t
- F
- Z
- Chi Square

Statistical tests should be selected for each MOE. Table 2-11 summarizes statistical test selection factors. Figure 2-11 shows the Step 7 flow. The selected test for each MOE should be entered on the sample Statistical Test listing form shown in Exhibit 2-7. Note that any test not specifically mentioned in this discussion should be justified and referenced on this form.

TABLE 2-11. STATISTICAL TEST SELECTION FACTORS

MOE	Data Type	Applicable Test(s)	Comments
Accidents	Rare Event	Poisson Distribution Test	Chi Square could also be used.
Brake Applications	Categorical	Chi Square	Nonparametric - A minimum of 5 data points.
Conflicts, Traffic	Categorical	Z Test	Parametric - Use with large samples ($N \geq 30$).
		Chi Square	Nonparametric - Use with small samples ($N \leq 30$).
Compliance	Categorical	Chi Square	See "Brake Applications."
Driving Slowly	Categorical	Chi Square	See "Brake Applications."
Delay	Continuous	t Test	Parametric
		Chi Square	Use if parametric assumptions cannot be met.
Encroachments	Categorical	Chi Square	See "Brake Applications."
Erratic Maneuvers, Gore Weaves, Last Minute Lane Changes	Categorical	Z Test Chi Square	See "Conflicts."
Head Turning Movements	Categorical	Chi Square	See "Brake Applications."
Lane Changes	Categorical	Chi Square	See "Brake Applications."

TABLE 2-11 (continued)

MOE	Data Type	Applicable Test(s)	Comments
Lateral Placement	Continuous	t Test	Use t (see "Delay") if normally distributed.
	Categorical	Chi Square	Use Chi Square if skewed or if data are scored categorically.
Merges	Categorical	Chi Square	See "Brake Applications."
Passes, Abortive	Categorical	Chi Square	See "Brake Applications."
Passing Distance	Continuous	t Test	See "Delay."
Passing Frequency	Categorical	Chi Square	See "Brake Applications."
Passing Time	Continuous	t Test	See "Delay."
Passing Type	Categorical	Chi Square	See "Brake Applications."
Points of Entry	Continuous or Categorical	t Test or Chi Square	See "Lateral Placement."
Speed, Spot	Continuous	t Test	See "Delay."
Speed Profile	Categorical	Chi Square	See "Brake Applications."
Time Headway	Categorical or Continuous	t Test or Chi Square	See "Lateral Placement."
Time Through Intersections	Continuous	t Test	See "Delay."
Variance: Speed, Lateral Placement, etc.	Continuous	F Test	Parametric - Tests for differences in variances of certain MOE's.

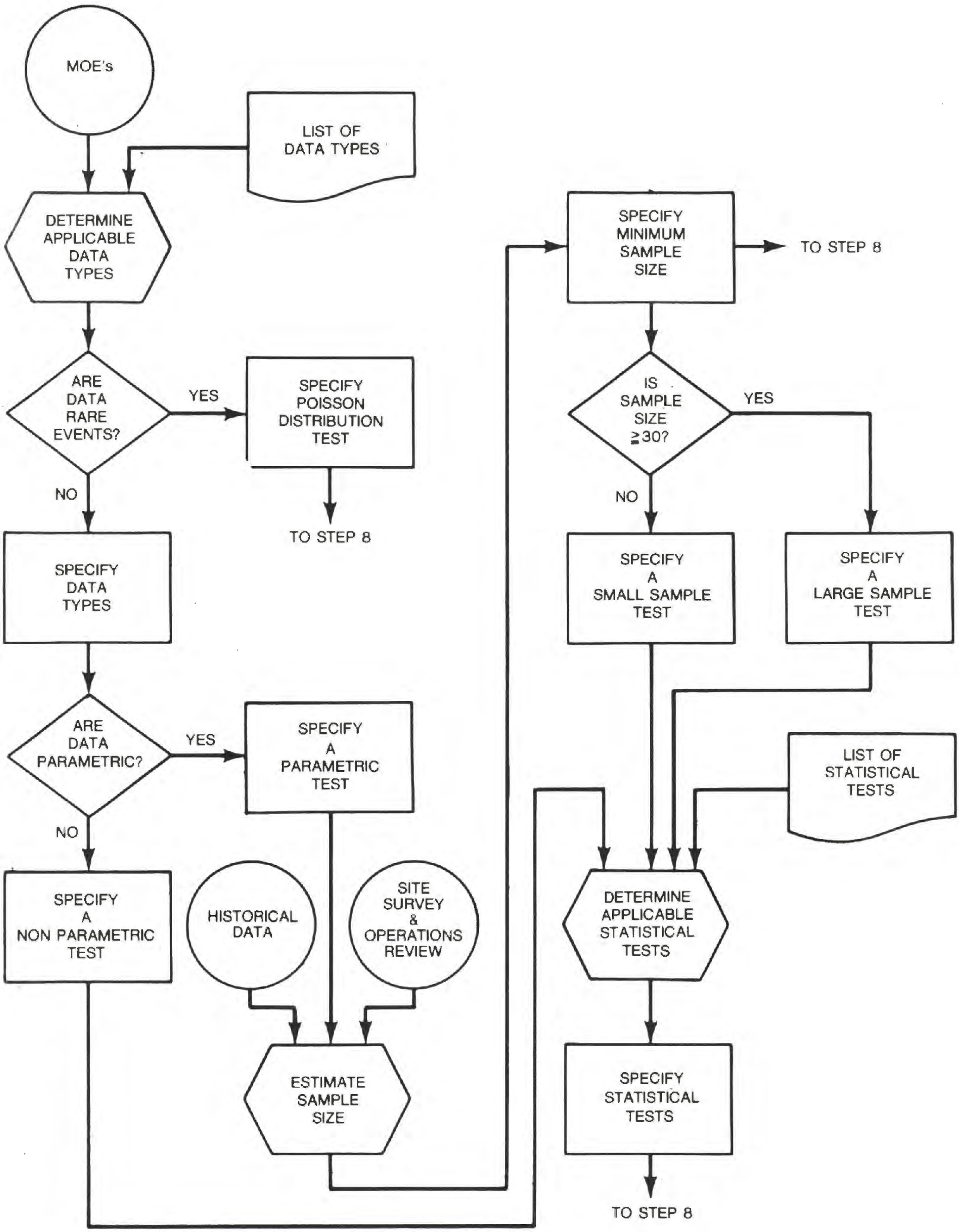


Figure 2-11.
Flow, Step 7 - Specify Statistical Tests

STATISTICAL TESTS—STEP 7

PROJECT: *Morris Ave. Interchange Lane Drop*

PROJECT NO.: *14-17*

EVALUATOR/DATE: *HL 9/18/80*

MOE: *Accidents*

TEST: *Poisson*

COMMENTS: *Have 3 years "Before" data in frequency on print-out–must break down by type–need volume.*

MOE: *Erratic Maneuvers*

TEST: *Z Test*

COMMENTS: *Collect volume.*

MOE: *Spot Speed*

TEST: *t Test*

COMMENTS:

Exhibit 2-7.
Sample Statistical Test Listing Form

STEP 8. Specify Confidence Level

Statistical significance is established on the basis of whether or not differences in the MOE's between the "Before" and "After" phases can be attributed to chance. Significance is expressed in terms of the confidence level, which is the degree of confidence, in percent, that the result was not due to chance. The probability, in decimals, of the results being due to chance is called the level of significance. It is an analogous term also often used to express statistical significance. All statistical tests are interpreted in terms of a confidence level.

Selecting an appropriate confidence level is not as clear cut as it might appear. While it may seem obvious that the way to assure that the results are not due to chance would be to specify a high value such that the chance probability is extremely low, doing so reduces the chances of the results being significant. On the other hand, setting the confidence level at a low value increases the likelihood of obtaining significance but also increases the possibility of the result actually being due to chance. There are two kinds of decision errors that can be made. The first is called a Type I error. It can occur when the confidence level is low and leads to a true chance result being deemed significant. The second is called a Type II error. It can occur when the confidence level is high and leads to a truly significant result being attributed to chance and deemed not significant. The implication of a Type I error for decisionmakers is that they may be implementing a program on the basis of an erroneous assumption, i.e., that an improvement is significant.

The level ultimately specified, therefore, represents a trade-off between the need to maximize the use of an effective improvement, while at the same time being assured that the improvement is, in fact, effective. Table 2-12 shows a range of significance values that are conventionally used, along with the odds of the results being due to chance. As a rule of thumb, low-cost, one-shot improvements generally are evaluated with a lower confidence level, while high cost projects and projects which may ultimately lead to a wide application of an improvement generally use a higher level. Research projects, by convention, usually specify a 95% confidence (.05 level of significance).

TABLE 2-12.
CONVENTIONAL LEVELS OF CONFIDENCE AND SIGNIFICANCE (TYPE I ERROR)

Confidence Level	Level of Significance	Odds of Significant Results Actually Being Due to Chance
80%	.20	One in Five
90%	.10	One in Ten
95%	.05	One in Twenty
99%	.01	One in One Hundred
99.9%	.001	One in One Thousand

The level selected generally depends on the cost of being wrong in the evaluation's conclusions as well as the following factors: (1) Agency policy; (2) the importance of the project to an agency's overall policy of hazard correction; (3) the safety criticality of the problem; and (4) the estimated cost of the improvement.

If an agency has a policy, it should be used. If not, then the confidence level could be determined on the basis of program importance, safety criticality, and project cost. For example, if most factors are moderate, an 80% confidence (.20 level) might be used; if several factors are high and no factor is very high, a 90% confidence (.10 level) may be specified; and if any factor is very high, a 95% confidence (.05 level) could be selected. Given the nature of most traffic operations and Positive Guidance projects, an 80% confidence (.20 level) will generally suffice if there are no program implications. There will hardly ever be any situation where a 99% confidence (.01 level) or greater would be specified.

Figure 2-12 shows the Step 8 flow. Exhibit 2-8 shows a form used to specify the confidence level. It is recommended that the justification for the level selected be included on this form.

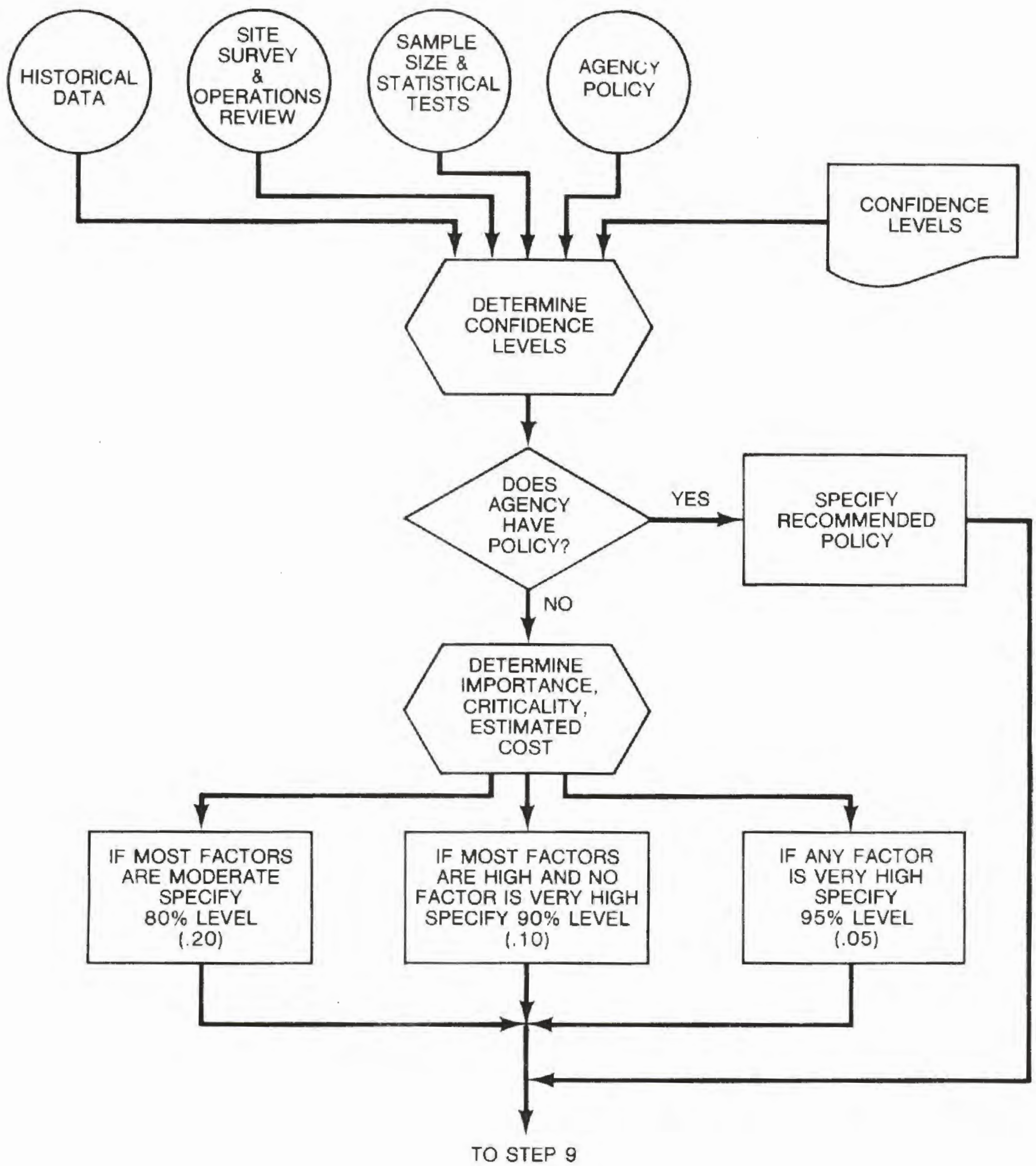


Figure 2-12.
Flow, Step 8 - Specify Confidence Level

CONFIDENCE LEVEL—STEP 8

PROJECT: *3rd Street RR Xing*

PROJECT NO.: *976 RR X*

EVALUATOR/DATE: *HL 7/6/79*

AGENCY POLICY: ☑ NONE ☐ SPECIFY ______

FACTORS:

PROGRAM IMPORTANCE:	☑ MODERATE	☐ HIGH	☐ VERY HIGH
SAFETY CRITICALITY:	☐ MODERATE	☑ HIGH	☐ VERY HIGH
PROJECT COST:	☐ MODERATE	☑ HIGH	☐ VERY HIGH

CONFIDENCE LEVEL
(LEVEL OF SIGNIFICANCE)

☐ 80% (.20)

☑ 90% (.10)

☐ 95% (.05)

☐ ______ Other (Specify)

COMMENTS: *Program implications if improvement shows long-term effectiveness.*

Exhibit 2-8.
Sample Confidence Level Form

STEP 9. Develop Sampling Plan

In Step 2, a target population of drivers and/or vehicles was identifed. The target population is all the drivers and/or vehicles that are involved with the problem at the site. Since it is usually not possible to collect data on an entire population, only a portion of the drivers and/or vehicles can be observed. The process used to select those drivers and/or vehicles for observation is called sampling, and the drivers and/or vehicles actually observed is called the sample. This step considers the two basic sampling issues: (1) Including a sufficient number of drivers and/or vehicles in the sample for descriptive purposes and to test for significance; and (2) drawing a sample that is representative of the target population.

All other things being equal, the larger the sample size, the more likely it is to be representative. Thus, sample size should be as large as possible. However, since it is not always feasible or desirable to spend too much time at a site, minimum sample size requirements have been specified. It may not be possible to achieve a large sample size, and grouping may be needed to achieve even a minimum sample. Specifying a minimum sample size often provides input into the selection of an appropriate sampling technique and frequently yields an estimate of how long a period must be spent in the field.

In specifying a minimum sample size, two factors must be determined, permitted error and confidence. Permitted error relates to how closely the obtained sampled value of the MOE must be to the true population parameter. For example, if spot speed is randomly sampled and an average speed is obtained, this value will probably differ from the average value that would be obtained if spot speeds were taken on every vehicle for a full year. How much the sampled value can differ is an engineering judgment, based on how accurate the sample must be for diagnostic purposes. Permitted error thus relates to sampling accuracy. It is desirable to have as accurate a sample as is practicable. However, the less sampling error that is permitted, the larger the sample must be. Similarly, the greater the confidence in the accuracy, the larger the sample size required.

Confidence levels were selected in Step 8. In this step, permitted error must be determined. There is no hard and fast rule for the selection of a permitted sampling error. If the evaluator can estimate the accuracy needs in specific terms, this value can be used in an existing sample size formula (see Box and Oppenlander, 1976), or the evaluator can choose a value in Table 2-13 close to the needed accuracy. If the evaluator cannot estimate accuracy needs, he or she can, as a rule of thumb, consider importance, criticality and cost in the same manner as Step 8. That is, if factors are moderate to high, a larger permitted error (lower accuracy) can be used, and if factors are high to very high, a smaller permitted error (higher accuracy) would be warranted.

Table 2-13 presents minimum sample size requirements for a range of MOE's. The table is structured for the three recommended confidence levels and two levels of sampling accuracy.

TABLE 2-13. MINIMUM SAMPLE SIZE

MOE	PERMITTED SAMPLING ERROR	MINIMUM SAMPLE SIZE CONFIDENCE LEVEL 80%	90%	95%
Continuous Data:				
Delay--Each Approach	5%	656	1,080	1,540
	10%	165	270	380
Speed--Each Observation*	±1.0 mph	41	68	96
	+5.0 mph	30**	30**	30**
Categorical Data: All Categorical	5%	30**	270	380
MOE's	10%	30**	70	100

*100 is recommended.
**Minimum sample size to meet statistical assumptions.

Samples are often drawn at <u>random</u> so that each <u>element</u> (i.e., vehicle, driver, condition, etc.) has an equal chance of being included. However, for purposes of the evaluation, it is likely that there will be elements of selection, referred to as <u>stratification.</u> In stratified sampling, the entire population of elements is segmented into homogeneous groups (strata) and each stratum is independently sampled. Since "Before" conditions are structured in terms of target conditions and/or groups (e.g., wet weather, night conditions, commuters, trucks, etc.), these attributes are used to stratify the samples. It may also be useful to employ a <u>systematic</u> sampling technique. With this technique, the target population is divided into <u>clusters</u> of equal elements (e.g., all vehicles in a given time period) and each n^{th} cluster sampled. It is acceptable to use a systematic technique with stratified random sampling.

The selection of a sampling plan is based on a number of considerations. If there are no specific target conditions or time pressures, then a purely random sampling technique could be used. That is, dates and times for data collection during a year are selected using a table of random numbers. This technique can be augmented by systematic sampling. Once the first date and time is selected, every n^{th} day and time is sampled until the sample size requirement is met.

Given target "Before" conditions and groups, and "real world" time pressures, a stratified sampling technique is recommended. The "Before" conditions (see Step 2) are used as the strata. Depending on the nature of the MOE's and target populations, data are collected on all attributes of interest until the desired sample size or greater is achieved. This may encompass all vehicles and/or maneuvers in the traffic stream, or only certain vehicles (e.g., trucks, RV's) or drivers (e.g., commuters, out-of-state license plates) or maneuvers (e.g., all passes or exit maneuvers). In locations with a high traffic volume or a large incidence of the MOE's of interest, it is often useful to employ a systematic, time-based sampling technique. For example, all vehicles of interest are sampled for one-half hour every 2 hours until the desired sample size is achieved. This technique is useful with time-lapse data collection.

Figure 2-13 presents the Step 9 flow. In Exhibit 2-9, an example of a form used to specify a minimum and desirable sample size and the sampling plan is shown.

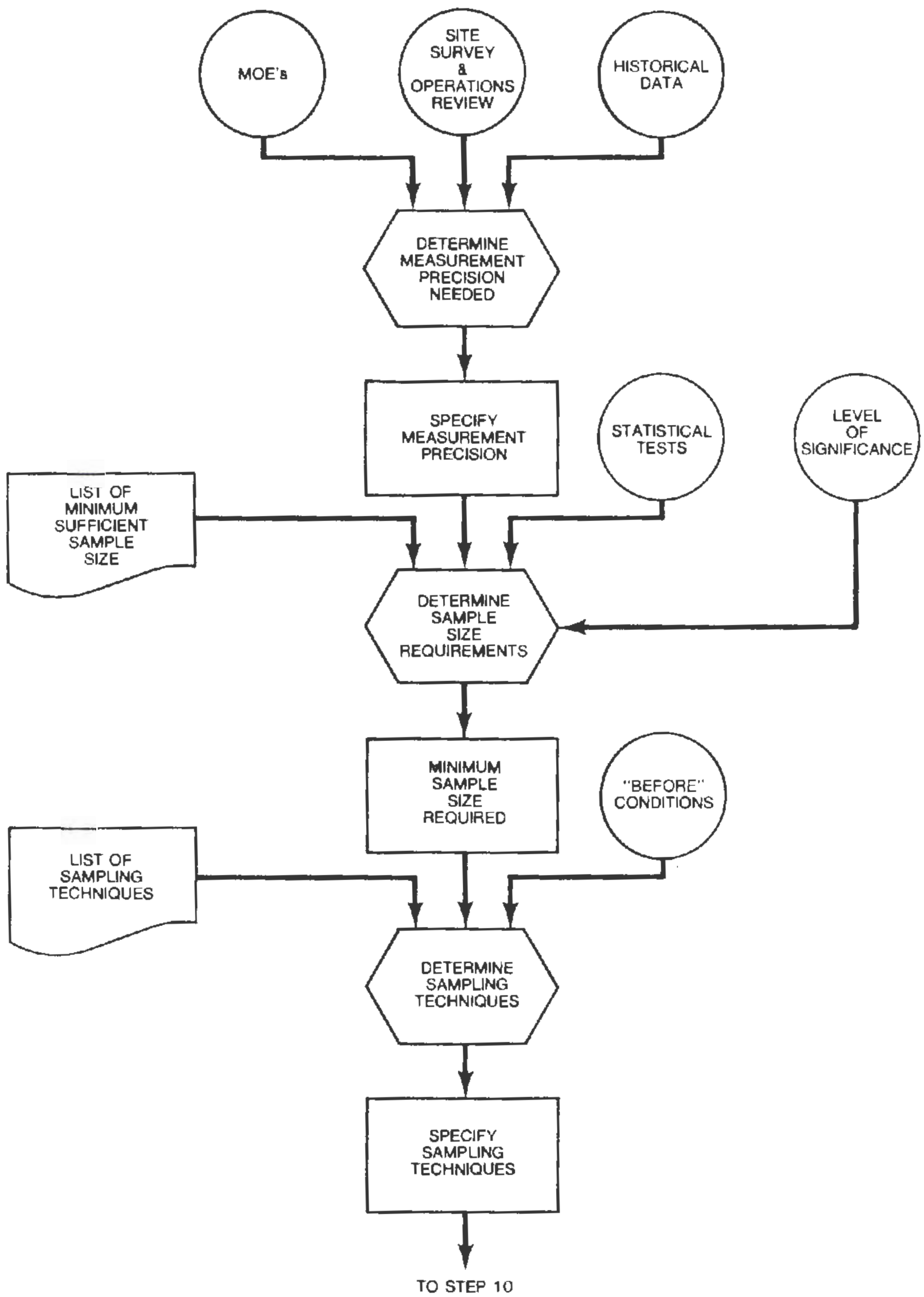


Figure 2-13.
Flow. Step 9 - Develop Sampling Plan

SAMPLING PLAN—STEP 9

PROJECT: *I 378–Lavell Road Interchange Lane Drop*

PROJECT NO.: *739–I 38B*

EVALUATOR/DATE: *HL 7/17/79*

MOE: *Erratic Maneuvers* PERMITTED ERROR: *10%*

MINIMUM SAMPLE SIZE: *70* DESIRED SAMPLE SIZE: *270*

SAMPLING PLAN: *Stratified for "Before" conditions–Systematic-time-based 1/2 hr on 1/2 off for 3 hours in p.m. peak for Tuesday, Wednesday, Thursday.*

MOE: *Speed* PERMITTED ERROR: *± 1 MPH*

MINIMUM SAMPLE SIZE: *30* DESIRED SAMPLE SIZE: *68*

SAMPLING PLAN: *S/A above–will try for 100 min–will stratify for out-of-town license plates and collect for 1st vehicles in platoons using systematic-time-based as above.*

MOE: *Accidents* PERMITTED ERROR: *N/A*

MINIMUM SAMPLE SIZE: *—* DESIRED SAMPLE SIZE: *—*

SAMPLING PLAN: *Will collect all accidents at site for 3-year period in "After"–will use historical accident records in "Before."*

Exhibit 2-9.
Sample Sampling Plan Listing Form

STEP 10. Prepare Data Collection Plan and Schedule

The final step in the Evaluation Plan development is to produce a detailed Data Collection Plan and Schedule. The plan should include applicable features from the project and each step in sufficient detail to serve as a working plan and schedule for "Before" data collection. The plan should also include a schedule for all aspects of the Improvement Development and Improvement Evaluation phases.

Primary emphasis in preparing the plan should be in specifying the following Pre-Data Collection ("Before" phase) details:

Equipment - All required equipment such as radar speed meters, counters, time-lapse cameras, etc., should be identified by type, make, model, etc. If the requisite equipment is available, it should be identified by serial number or other identifying information. If not available, it should be scheduled to be procured in sufficient time to be available for "Before" data collection. Equipment calibration requirements should be noted, and all calibration and testing should be scheduled to insure proper operation. Consideration should be given to providing spare or backup units in the event of equipment failure. If more than one of the same unit is being used, a comparability check should be scheduled to assure data comparability or to enable adjustments to be made. It is very important to keep a record of all equipment used to insure comparability by using the same or similar equipment throughout the evaluation.

Personnel - Personnel requirements for field data collection should be specified and scheduled. Personnel should be identified by function and name (if possible) to determine their availability both for the "Before" and "After" phases. It is best to use the same personnel throughout, particularly when there is manual data collection. If personnel are not members of the staff, individuals needed to perform required functions must be recruited. Care should be taken in recruiting data collection personnel, particularly if they are to be hired on a casual or project-specific basis (such as summer help or college students). Accuracy checks should be considered for personnel. It is important to determine training needs. Training may be necessitated when using unfamiliar, new or casual personnel, or when using new equipment or procedures. Training and, in some cases, retraining may be needed if teams are used, if judgments must be made under time pressures, or if unfamiliar forms are used. In any case, training needs should be determined, training procedures dealing with all aspects of the data collection task (e.g., equipment, measures, procedures, forms, contingencies, etc.) structured, and training scheduled.

Data Collection Forms - All data collection form requirements should be formulated prior to data collection. If forms are not available, they should be scheduled for design and production. They should be scheduled for pretest prior to production to make sure that they are readily usable in the field and that ample space has been provided for data entry. One form that should always be provided is a daily log such as the one shown in Exhibit 2-10. Its purpose is to record the phase, dates and times, project data collection personnel, conditions, events, and any deviations from the Evaluation Plan.

Coordination - Two types of coordination may be required. The first is coordination within project staff. This coordination is very important when one or more teams are collecting data, or when more than one shift is involved. A second kind of coordination is between project personnel and authorities such as police, other agencies, etc. In both cases, coordination requirements should be formulated and coordination scheduled.

Procedures - Prior to going into the field, a set of instructions should be prepared, indicating all aspects of data collection. Included in these instructions should be equipment and personnel location and how, when and where the data are to be collected. Procedures for crew shifts (if applicable) should be developed and procedures for how to handle contingencies, e.g., when to cancel data collection due to accidents, rain, snow, etc., should be clearly defined.

Pilot Testing - One way to assure that all aspects of the data collection phase will go smoothly is to design and schedule pilot tests. Pilot or pretesting prior to the start of the "Before" data collection phase may be warranted: When unfamiliar or new equipment is involved; when there is a need to verify equipment calibration; when unfamiliar or new personnel are used; when training is required; when unfamiliar data collection forms are used; when there is a need to practice on-site coordination; when unfamiliar or complex procedures, equipment location or personnel location is used; when there is a need to coordinate data collection team activities such as shift changes; and when data collection may be obtrusive. If pilot testing is required, details should be specified, and tests scheduled. This will enable modifications to be made to the "Before" data collection plan, and will allow the plan to be finalized.

Figure 2-14 provides the Step 10 flow. A sample plan is shown in Exhibit 2-11.

DAILY LOG	
PROJECT: *I 27 & US 329 Split*	
PROJECT NO.: *329 B*	
EVALUATOR/DATE: *HL 8/6/79*	
PHASE: ☑ "BEFORE" (STEP 12) ☐ IMPLEMENTATION AND ACCLIMATION (STEP 15) ☐ "AFTER" (STEP 17)	
DATA COLLECTION PERSONNEL: *R. Smith & J.J. Kospert*	
DATE/TIME	CONDITIONS, EVENTS OR DEVIATIONS
8/5/79	
8:00 a.m.	*Set up equipment–day clear, traffic light*
9:00 a.m.	*Begin filming–Roll #1*
9:15 a.m.	*Film broke–Data session stopped & rescheduled for 9:30 a.m.*
9:30 a.m.	*Re-start–Roll #2*
11:00 a.m.	*Accident–Rescheduled p.m. start to 3:30 p.m.*
3:30	*Re-start–Roll #3*

Exhibit 2-10.
Sample Daily Log Form

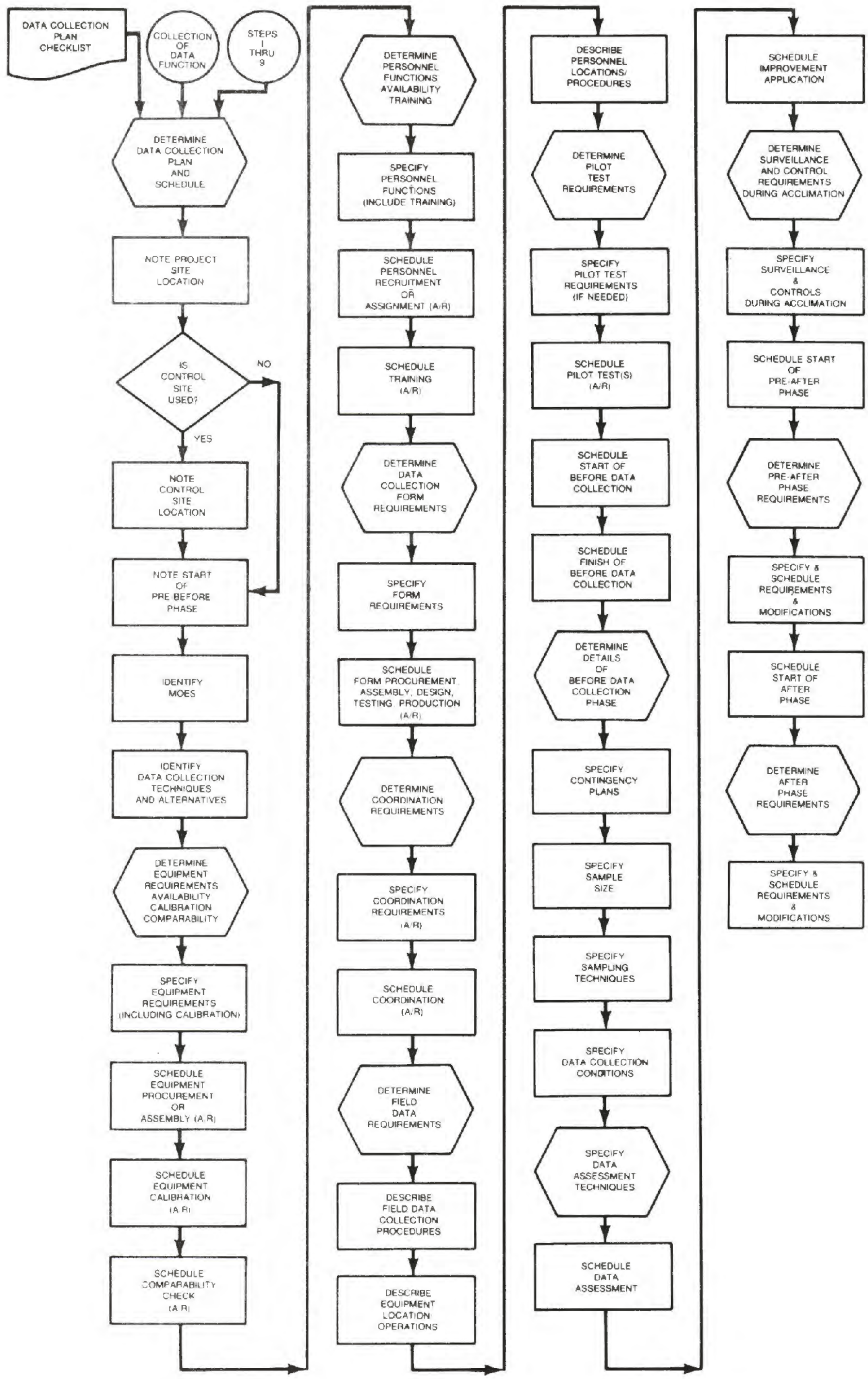


Figure 2-14.
Flow, Step 10 - Prepare Data Collection Plan and Schedule

DATA COLLECTION PLAN AND SCHEDULE—STEP 10

PROJECT: *Interchange of US 43 & DeKalb St.*

PROJECT NO.: *127-46*

EVALUATOR/DATE: *HL 6/6/79*

PROJECT LOCATION: *US 43 & DeKalb St.–Sulph City*

CONTROL SITE LOCATION: *None*

DATE FOR:

START OF PROJECT: *5/4/79*

START OF PRE-DATA COLLECTION "BEFORE" PHASE: *7/17/79*

MOE's:

1— *Accidents*

2— *Erratic Maneuvers*

3— *Spot Speeds*

4— ______

DATA COLLECTION TECHNIQUE(S):

MOE: *Accidents*
PRIMARY: *Historical Records*
ALTERNATIVE: *None*

MOE: *Erratic Maneuvers*
PRIMARY: *Time Lapse*
ALTERNATIVE: *Manual Observation*

MOE: *Spot Speed*
PRIMARY: *Radar*
ALTERNATIVE: *Electronic Stop Watch*

MOE: ______
PRIMARY: ______
ALTERNATIVE: ______

Exhibit 2-11.
Sample Data Collection Plan and Schedule Form

DATA COLLECTION PLAN AND SCHEDULE	**PROJECT NO.** *127-46*

EQUIPMENT REQUIREMENTS:
(NOTE TYPE, QUANTITY,
AVAILABILITY, SERIAL NO.)

MOE: *Accidents*

TECHNIQUE: *Historical Records*

State computer print-out.

DATES FOR PROCUREMENT *N/A* OR ASSEMBLY *N/A*

MOE: *Erratic Maneuvers*

TECHNIQUE: *Time Lapse*

Kodak 123–1 required & available in stock room S/N 127Z96

DATES FOR PROCUREMENT ______ OR ASSEMBLY *7/17/79*

MOE: *Spot Speed*

TECHNIQUE: *Radar Speed Meter*

XYZ radar speed meter–2 required–one available in stock room–S/N 3334.
One must be procured.

DATES FOR PROCUREMENT *6/9/79* OR ASSEMBLY *7/17/79*

MOE: *Spot Speed*

TECHNIQUE: *Electronic Stop Watch*

ABC stopwatches 2 required–available in stock room–S/N ZZ1 & ZZ7

DATES FOR PROCUREMENT ______ OR ASSEMBLY *7/17/79*

Exhibit 2-11.
(Continued)

DATA COLLECTION PLAN AND SCHEDULE	PROJECT NO. *127-46*

CALIBRATION AND COMPARABILITY CHECKS:

EQUIPMENT: *XYZ Speed Meters*
SERIAL NO.: *SN 3334 + New Procurement*
REQUIREMENT: *Calibration & Comparability Check*

EQUIPMENT: *ABC Stop Watch*
SERIAL NO.: *ZZ1 & ZZ7*
REQUIREMENT: *Calibration & Comparability Check*

EQUIPMENT: ______
SERIAL NO.: ______
REQUIREMENT: ______

EQUIPMENT: ______
SERIAL NO.: ______
REQUIREMENT: ______

DATES FOR CALIBRATION: *7/18/79* OR COMBATIBILITY CHECK: *7/19/79*

PERSONNEL REQUIREMENTS (NOTE FUNCTIONS, INDIVIDUALS, AVAILABILITY)

Two two-person teams needed
Both data collectors
Team A–R.L. Larson + assistant
Team B–A.B. Sundon + assistant
Larson & Sundon on board
Two "casuals" will be recruited

DATES FOR RECRUITMENT: *7/12/79* OR ASSIGNMENT: *7/12/79*

Exhibit 2-11.
(Continued)

DATA COLLECTION PLAN AND SCHEDULE | **PROJECT NO.** 127-46

TRAINING REQUIREMENTS:

New personnel must be trained in the use of radar speed meters & electronic stop watches.

DATES FOR TRAINING: *8/13-8/15-79*

FORM REQUIREMENTS (NOTE FORMS, AVAILABILITY):

Daily log forms–available.
Need form for speed monitoring–will use form in manual of Traffic Engineering Studies.

DATES FOR DESIGN: 7/31/79
TESTING:
PRODUCTION: 8/3/79

DATES FOR PROCUREMENT: *8/13/79* OR ASSEMBLY: *8/13/79*

COORDINATION (DESCRIBE REQUIREMENTS):

Teams must coordinate crew shift.
Must coordinate with State HP.

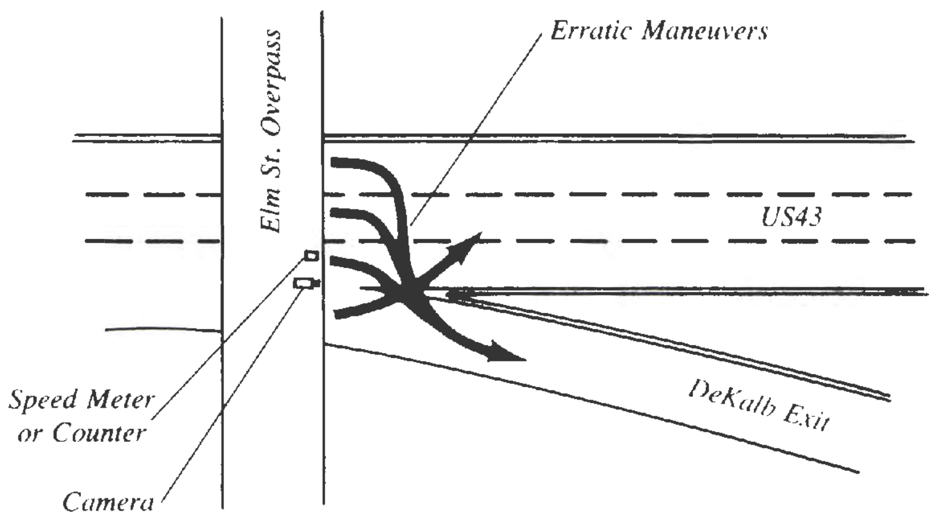
DATES FOR COORDINATION: *Meet with police 7/23/79*

Exhibit 2-11.
(Continued)

DATA COLLECTION PLAN AND SCHEDULE	PROJECT NO. 127-46

DATA COLLECTION PROCEDURES, EQUIPMENT LOCATION AND OPERATION, AND PERSONNEL LOCATION AND PROCEDURES, OPERATIONAL DEFINITIONS:

MOE: *Erratic Maneuvers, Spot Speeds*



Crews to set-up 1 hour prior to data collection.
Speed meters to be covered up.
Crew A to monitor AM peak.
Crew B to monitor PM peak.

Erratic maneuvers to be recorded on film & back-up manual observation.
Monitor CB 19 for "Fuzz Buster" chatter.

Erratic maneuvers as in diagram–in vicinity of physical & painted gore.
Spot speeds for vehicles in curb lane.

Exhibit 2-11.
(Continued)

DATA COLLECTION PLAN AND SCHEDULE	PROJECT NO. 127-46

PILOT TESTING CHECK LIST:

(1) NEW OR UNFAMILIAR: EQUIPMENT	☑ YES	☐ NO
DATA COLLECTION TECHNIQUES	☐ YES	☑ NO
PERSONNEL	☑ YES	☐ NO
FORMS	☑ YES	☐ NO
(2) IS TRAINING REQUIRED?	☑ YES	☐ NO
(3) IS DATA COLLECTION COMPLEX?	☑ YES	☐ NO
(4) IS COORDINATION NEEDED?	☑ YES	☐ NO
(5) IS OBTRUSIVENESS A FACTOR?	☑ YES	☐ NO
(6) OTHERS ______ (SPECIFY)	☐ YES	☐ NO

PILOT TESTING (INDICATE DETAILS)

- *Crews & equipment checked out two days prior to field data collection – develop and look at film*
- *Check if crew can be observed or if radar meters are being picked up*
- *Rehearse crew shift*
- *Check casuals for training and accuracy*
- *Filming position etc.*

DATE(S) FOR PILOT TESTING	START 8/20/79
	FINISH 8/21/79

Exhibit 2-11.
(Continued)

DATA COLLECTION PLAN AND SCHEDULES	PROJECT NO. 127-46

DATE FOR START OF "BEFORE" DATA COLLECTION: *8/27/79*

DATA COLLECTION CONDITIONS (TIME OF DAY, DAY OF WEEK, ENVIRONMENTAL, ETC.)

- *A.M. Peak 8:30-10:30*
  P.M. Peak 4:30-6:30
- *Tues, Wed, Thurs*
- *Days should be clear, no rain*
- *Daylight conditions*

CONTINGENCY PLANS:

Terminate if rain or accidents occur – can use following week

Use manual counts if camera or speed meters fail

Exhibit 2-11.
(Continued)

DATA COLLECTION PLAN AND SCHEDULE	PROJECT NO. *127-46*

SAMPLE SIZE(S):

MOE: *Accidents*

MINIMUM SAMPLE: *All*

MOE: *Erratic Maneuvers*

MINIMUM SAMPLE: *70-270 Desirable*

MOE: *Spot Speeds*

MINIMUM SAMPLE: *100*

SAMPLING TECHNIQUE(S): MOE: *Accidents –*

Historical Records

MOE: *Erratic Maneuvers*

(For film) Systematic time based – 1/2 hr. "on" – 1/2 hr. "off"
(100% if manual)

MOE: *Spot Speeds*

Will collect 100% for curb lane speed –
only lead vehicles in platoon.

DATA ASSESSMENT (QUANTITY, QUALITY)

Check speeds each day
Reduce film prior to pulling down field set-up

DATE(S) FOR DATA ASSESSMENT:

9/10-9/13/79

DATE FOR FINISH OF "BEFORE" DATA COLLECTION: *9/15/79*

Exhibit 2-11.
(Continued)

DATA COLLECTION PLAN AND SCHEDULE	PROJECT NO. *127-46*
DATE FOR APPLICATION OF IMPROVEMENT:	*Late Nov., early Dec. 1979*

SURVEILLANCE AND CONTROLS DURING IMPLEMENTATION AND ACCLIMATION PERIOD:

- *Continue coordination with police*
- *Visually observe operations*
- *Monitor volume*
- *Monitor accidents*

DATE FOR START OF PRE-DATA COLLECTION—"AFTER" PHASE

June 2, 1980

PILOT TESTING, TRAINING, ETC. (INDICATE DETAILS)

Retrain personnel
Train new personnel
Repeat pilot tests

DATE(S) FOR PILOT TESTING, TRAINING: START *June 8, 1980*
FINISH *June 10, 1980*

DATE FOR START OF "AFTER" DATA COLLECTION:

July 8, 1980

DATES FOR DATA ASSESSMENT: *July 10, 1980*

DATE FOR FINISH OF "AFTER" DATA COLLECTION:

July 12, 1980

DATES FOR REDUCTION OF DATA, ASSESSMENT OF RESULTS:

July 12, 1980

TARGET DATE FOR FINAL REPORT:

Late 1980

Exhibit 2-11.
(Continued)

COLLECTION OF EVALUATION DATA

The purpose of this section is to bring together and briefly describe the various data collection activities of the evaluation. The emphasis in this section is on applying the Evaluation Plan.

Because of the nature of the evaluation process, i.e., data collected prior to improvement development is compared with data collected after the improvement is implemented to determine effectiveness, the activities that comprise the Collection of Evaluation Data phase are not performed consecutively. In terms of the project development cycle, Collection of Evaluation Data begins following the development of the Evaluation Plan. After the improvement has been developed, the Collection of Evaluation Data phase resumes with the application of the improvement and the collection of "After" data.

Overview

Figure 3-1 presents a functional flow of the steps used to collect evaluation data. Conceptually, the phase is structured in terms of "pre-implementation" where "Before" data are collected, "implementation" where the improvement is developed and applied, and "post-implementation" where "After" data are collected. This section is structured to enable personnel to apply each step sequentially. To maintain continuity, the steps of the Evaluation Data Collection phase have been numbered consecutively with those of the Evaluation Plan. The first step in this phase begins with Step 11.

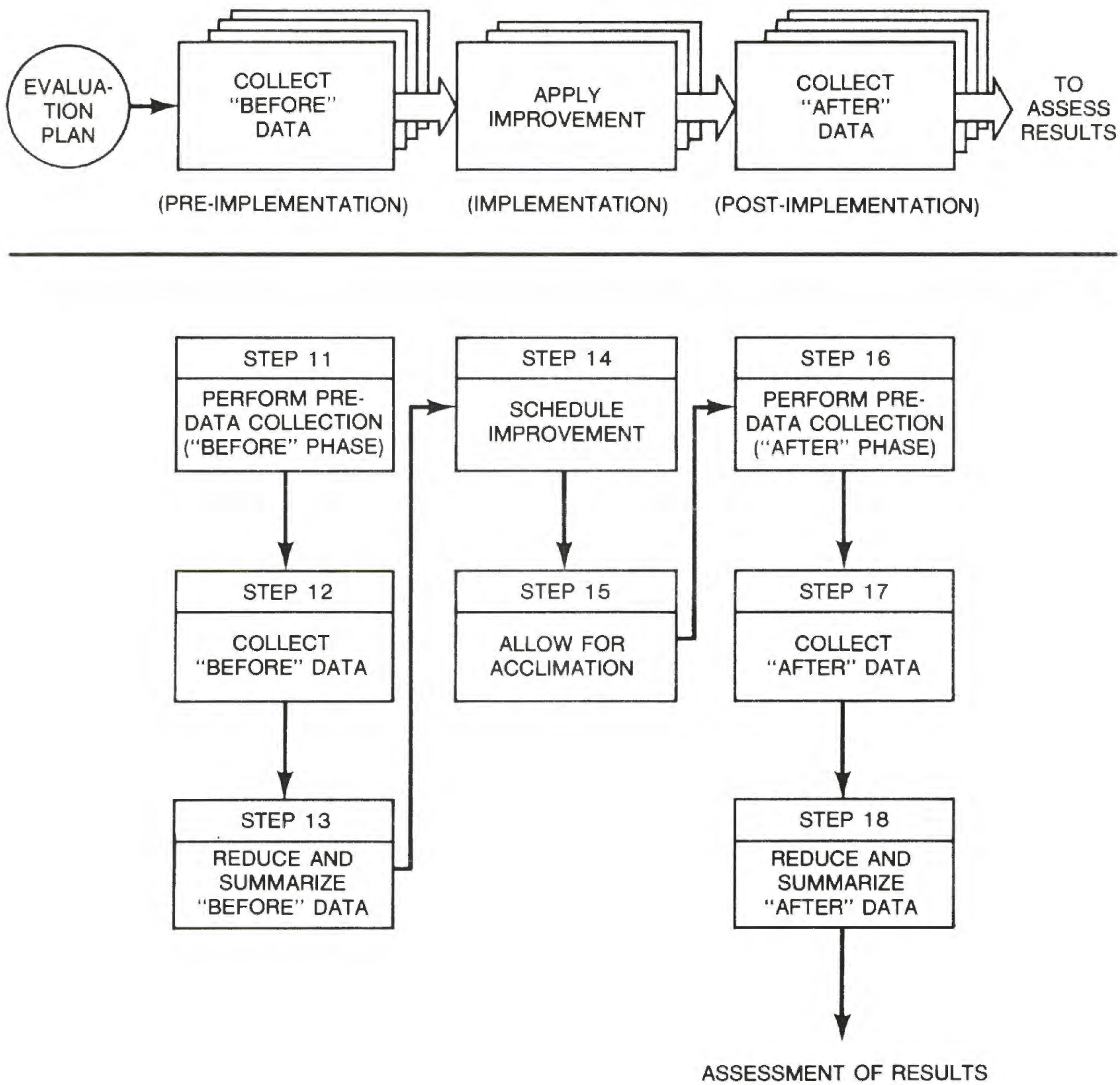


Figure 3-1.
Data Collection - Functional Flow

STEP 11. Perform Pre-Data Collection ("Before" Phase)

The logistical and operational matters that were identified in Step 10 for Pre-Data Collection ("Before" Phase) must be completed prior to collecting "Before" data. Performance of these activities will insure that data collection efforts run smoothly, and that the data gathered will be valid, reliable, meaningful and usable.

Equipment - All necessary equipment procurement, assembly, calibration, etc., should be performed.

Personnel - Personnel recruiting, scheduling and training, as required, should be accomplished.

Forms - Form needs should be taken care of in this step, including any necessary design and pretesting.

Coordination - Required coordination with all authorities should be instituted.

Procedures - Procedures should be finalized.

Pilot Testing - All identified pilot testing should be conducted during this step. Any changes brought about as a result of the pilot tests should be documented, and the "Before" Data Collection Plan revised.

Figure 3-2 presents the Step 11 flow. A Pre-Data Collection form is shown in Exhibit 3-1.

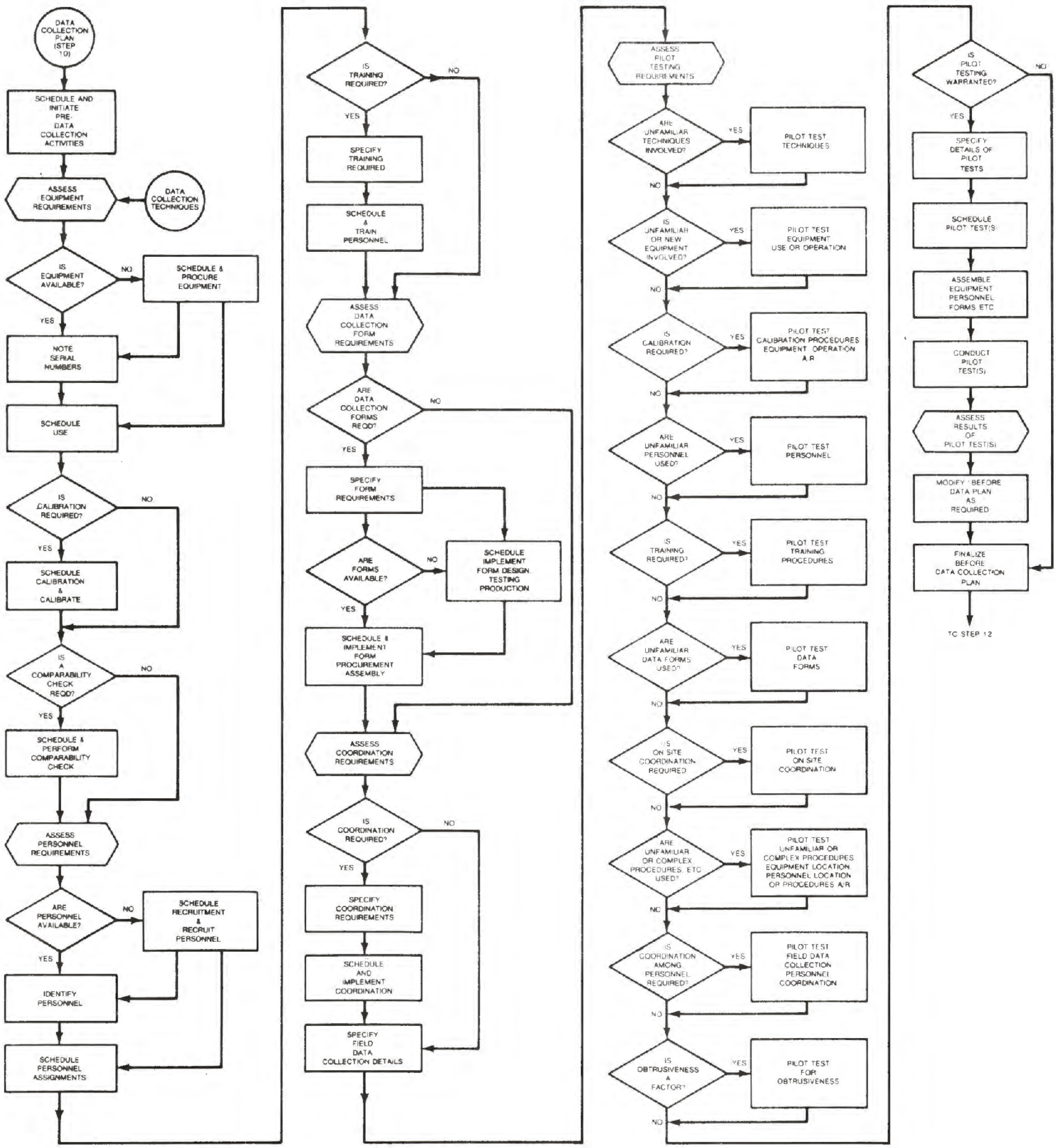


Figure 3-2.
Flow, Step 11 - Perform Pre-Data Collection ("Before" Phase)

PRE-DATA COLLECTION ACTIVITIES

PROJECT: *Interchange of US 43 & DeKalb Street*

PROJECT NO.: *127-46*

EVALUATOR/DATE: *HL 8/21/79*

PHASE: ☑ "BEFORE" (STEP 11) ☐ "AFTER" (STEP 16)

EQUIPMENT:	☑ PROCUREMENT	*7/17/79 –HL*
	☑ ASSEMBLY	*7/17/79 –HL*
	☑ CALIBRATION	*7/18/79 –HL*
	☑ COMPARABILITY CHECK	*7/19/79 –HL*
PERSONNEL:	☑ RECRUITMENT	*7/12/79 –HL*
	☑ ASSIGNMENT	*7/12/79 –HL*
	☑ TRAINING	*8/14/79 –HL*
FORMS:	☑ PROCUREMENT	*8/13/79 –HL*
	☑ ASSEMBLY	*8/13/79 –HL*
	☑ DESIGN, TESTING	*7/31/79 –HL*
	☑ PRODUCTION	*8/5/79 –HL*
	☑ COORDINATION	*7/27/79 –HL*
	☑ PROCEDURES	*8/20/79 –HL*
	☑ PILOT TESTING	*8/20-8/21/79 –HL*

PILOT TEST RESULTS:

Everything checked out except radar meters were picked up by trucks.

DATA COLLECTION PLAN CHANGES:

Delete radar speed meters from plan.

Exhibit 3-1.
Sample Pre-Data Collection Activities Listing Form

STEP 12. Collect "Before" Data

The "Before" data collection activities should begin once it has been established that all aspects of the Data Collection Plan have been accounted for, and that equipment, personnel and procedures are functioning properly. The reader should refer to Report FHWA-TO-80-2 for a discussion of field data collection considerations.

Data collection should be as inconspicuous as possible. Visible evidence of data gathering activities such as unusual vehicles, equipment, pneumatic tubes, obtrusive personnel, and radar may modify driver behavior. This can confound the MOE's being collected. One way to determine whether data gathering efforts are obtrusive is to monitor CB Channel 19 to see whether truckers have spotted the equipment and/or personnel. If there is any evidence that traffic is being affected, the activities should be stopped and procedures modified.

Since it will not be possible to replicate the "Before" conditions after an improvement has been applied, all aspects of the data collection should be constantly monitored to assure that the data being collected are properly gathered and both its quantity and quality are acceptable. For example, time-lapse film, if used, should be developed and viewed. The sampling plan and sampling techniques must be adhered to to insure representativeness and sample size suitability.

In the course of data collection, personnel should document all "Before" conditions in the daily log to assure that they are in accordance with the conditions specified by the Evaluation Plan. All deviations and use of contingency plans should also be documented. It is recommended that the daily log be kept current each day. This will help insure that "After" conditions are comparable and will be useful in data assessment.

Figure 3-3 shows the Step 12 flow. Daily log forms were presented in Step 10 (see Exhibit 2-10). Data collection forms for Step 12 (and Step 17) generally are developed on a project-specific basis. It is recommended that these forms include the same information contained in the heading on all forms shown herein and that the step number is identified on the form.

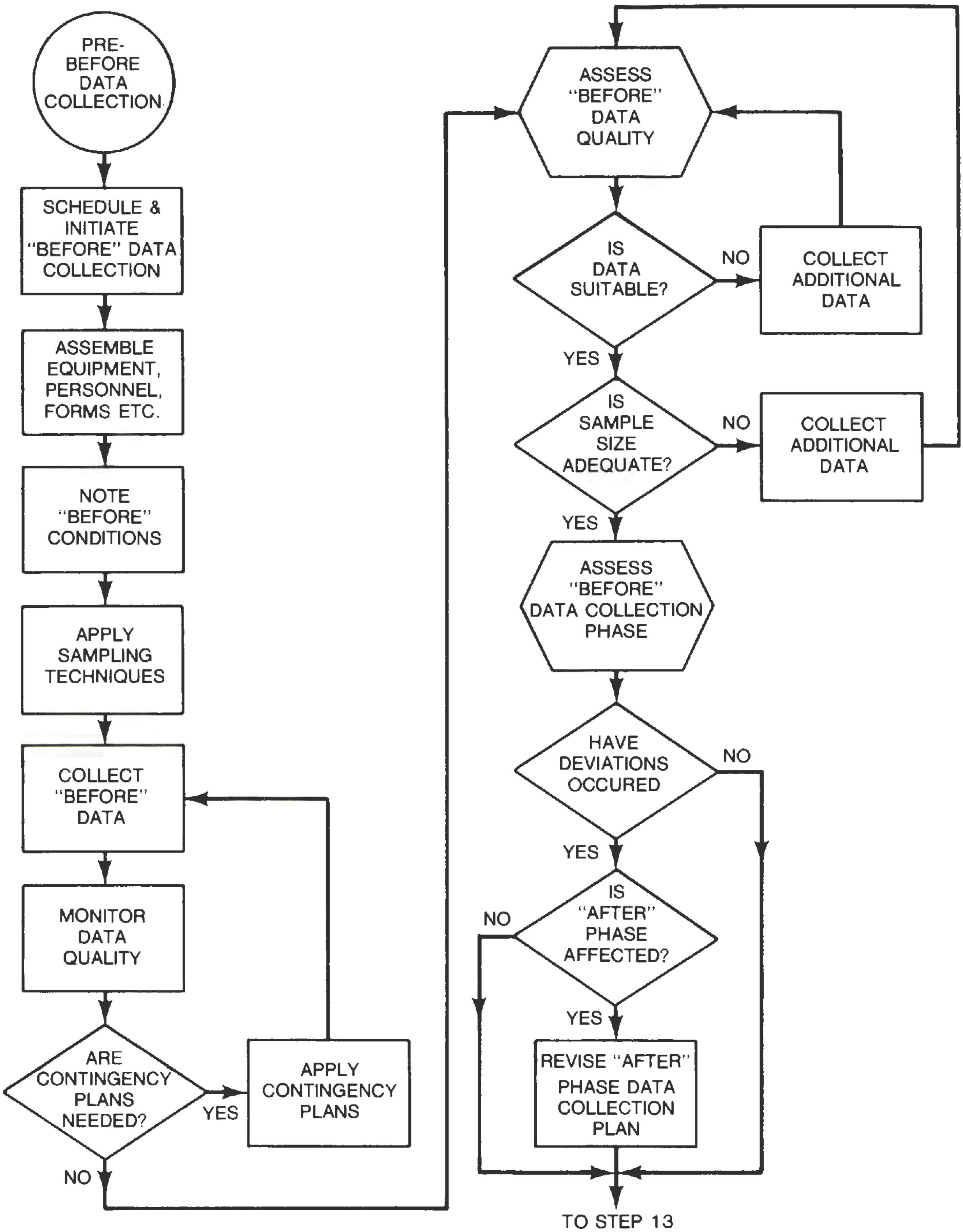


Figure 3-3.
Flow, Step 12 - Collect "Before" Data

STEP 13. Reduce and Summarize "Before" Data

The next step is to reduce and summarize the "Before" data. This will require the use of descriptive statistics, since the "Before" data are used to describe and diagnose problems at the site. The output of Step 13 will be summaries of applicable accidents, traffic performance (including individual driver performance), and/or systems performance MOE's. Figure 3-4 shows the Step 13 flow.

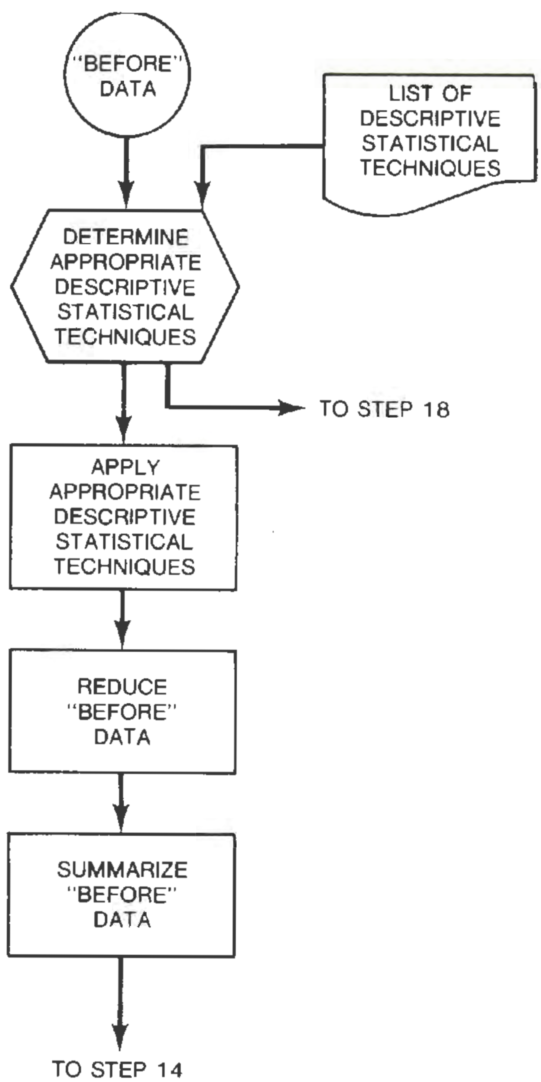


Figure 3-4.
Flow, Step 13 - Reduce and Summarize "Before" Data

Descriptive statistics generally are displayed as frequency distributions and arithmetic summations. It is often useful to show the shape of the distribution by presenting the data graphically, particularly if the information is to be presented to the general public, or when a formal report will be published and distributed.

Most MOE's can be summarized and displayed as frequency distributions. The exact form that a distribution takes depends on the nature of the data and the number of categories and conditions for which data were collected. Accidents, categorical data and continuous data can all be grouped and displayed in a contingency table in terms of the number, rate, and/or percentage of individuals or vehicles observed. In the case of continuous MOE's, the data must first be grouped before it can be put into an appropriate contingency table for display as a frequency distribution.

Rare Event Data - When accident MOE's are used, the data should be tabulated in an Accident Summary such as the one shown in Exhibit 3-2. The summary should also include exposure, since volume data will be required in Step 19 of the evaluation. Following completion of the summary, these data can then be displayed in terms of magnitude, percentage of total, or both as shown in the example in Exhibit 3-3.

Categorical Data - All categorical data is suitable for tabulation in a contingency table. This is true even for situations when there are only two categories, i.e., when a portion of the sample "does" or "does not" exhibit a particular behavior such as a "last minute lane change" or a "centerline encroachment." In these cases, indicating the proportion of the sample that exhibits the behavior is all that is needed. For example, indicating 5 percent encroachments implies 95 percent "no encroachments," and it is not necessary to develop a contingency table until the "Before" and "After" data are compared in Step 19. For multiple categories, a frequency distribution is warranted. Exhibit 3-4 shows a form for developing a contingency table for categorical data. Exhibit 3-5 shows a typical frequency distribution for the example contained in Exhibit 3-4.

ACCIDENT SUMMARY

PROJECT: *Intersection Project (Case Study 4)*

PROJECT NO.: *CS-4 (Site #1)*

EVALUATOR/DATE: *HL 3/20/80*

DATA SOURCE: *Datta, et al.*

TIME PERIOD: *1/71* TO *12/73*

LOCATION: ☑ PROJECT SITE ☐ CONTROL SITE

☑ "BEFORE"—STEP 13 ☐ "AFTER"—STEP 16

Accident Category	Total Accidents	Fatal Acc.	Fatalities	Injury Acc.	Injuries	PDO Acc.	Invol.
Surface Condition							
Dry	100	1	1	50	60	49	89
Wet	40	—		8	13	33	61
Snowy/Icy	—	—		—	—	—	—
Other	—	—	—	—	—	—	—
Total	141	1	1	58	73	82	150
Accident Type							
Overturn Collision with:							
Motor Veh.	136	—	—	54	69	82	150
Pedestrian	4	1	1	3	3	—	—
Pedal Cycle	1	—		1	1	—	
Animal	—	—	—	—	—	—	—
Fixed Object	—	—	—	—	—	—	—
Other	—	—	—	—	—	—	—
Total	141	1	1	58	73	82	150
Two Veh. Accidents							
Opposite Direction	10	—	—	8	13	2	4
Same Direction	40	—	—	4	8	36	70
One Veh. Stopped	5	—	—	1	1	4	7
One Veh. Entering Ramp	32	—	—	16	16	16	32
One Veh. Exiting Ramp	40	—	—	25	30	15	27
Other	9	—	—	1	1	8	10
Total	136	—	—	54	69	82	150
Two Veh. Accident Types							
Head-on	10	—	—	8	13	2	4
Rear-end	45	—	—	5	9	40	77
Sideswipe	9	—	—	1	1	9	10
Angle	72	—	—	41	46	31	59
Other	—	—	—	—	—	—	—
Total	136	—	—	54	69	82	150

Exhibit 3-2.
Sample Accident Summary and Exposure Listing Form

ACCIDENT SUMMARY

EXPOSURE *(all 5 sites)*

Site	Project* Length	Length of Time Period	AADT	Exposure Veh. X or Veh. Mi. ___
1. 1971	N/A	1. 365 days	1. 7,000	2.56×10^6
1972		365 days	7,100	2.59×10^6
1973		365 days	7,200	2.63×10^6
Subtotal		—	—	7.78×10^6
2. 1971	N/A	2. 365 days	2. 6,500	2.37×10^6
1972		365 days	6,800	2.48×10^6
1973		365 days	7,100	2.59×10^6
Subtotal		—	—	7.54×10^6
3. 1971	N/A	3. 365 days	3. 7,300	2.66×10^6
1972		365 days	7,300	2.66×10^6
1973		365 days	7,500	2.74×10^6
Subtotal		—	—	8.06×10^6
4. 1971	N/A	4. 365 days	4. 8,000	2.92×10^6
1972		365 days	8,400	3.06×10^6
1973		365 days	8,600	3.14×10^6
Subtotal		—	—	9.12×10^6
5. 1971	N/A	5. 365 days	5. 7.300	2.66×10^6
1972		365 days	7,500	2.66×10^6
1973		365 days	7,700	2.81×10^6
Subtotal		—	—	8.21×10^6
Total	—	—	—	40.71×10^6

*For vehicle-mile units of exposure (only)

Exhibit 3-2.
(Continued)

FREQUENCY DISTRIBUTION FOR "POOKS" HILL ACCIDENTS					
Distribution of Accidents (%)					
Year	Accident Frequency	Side-Swipe	Fixed Object	Rear End	Single Vehicle
1973	49	25%	49%	24%	2%
1974	38	16%	47%	24%	13%
1975	37	35%	41%	24%	0%
1976*	18	43%	33%	23%	0%
	*Partial				
TOTAL (Σ)	142				

Exhibit 3-3.
Example of a Frequency Distribution for Rare Event Data

Continuous Data - While continuous data are generally summarized as arithmetic summaries, it is often useful to develop frequency distributions to display interval MOE's such as speed. In fact, the data reduction task used to calculate means and standard deviations generates a frequency distribution which can be used as a data summary in this step. Exhibit 3-6 shows a form that can be used to generate frequency distributions. Its primary use is to develop arithmetic summaries and graphic displays for spot speed data. The example was taken from Box and Oppenlander (1976)**. In developing frequency distributions, it is necessary to first group the data into suitable classes. The class size should be such that it is neither too large nor too small, or that there are not too few or too many classes. Generally, from 8 to 20 classes is appropriate. To determine class size, the range is first determined by subtracting the lowest value in the data set from the highest value. In this case, the lowest speed is 30 mph and the highest speed is 53 mph, yielding a range of 23. The range is then divided by 8 and 20, yielding a maximum class size of 2.9 to a minimum class size of 1.15. A convenient class size is determined within these limits, in this example 2.0 is set. The limits are then converted to class as shown in the class limits column in Exhibit 3-6. The number of cases is then inserted in the "class frequency" column to complete the frequency distribution.

**This material is utilized by permission from the Institute of Transportation Engineers, 525 School Street SW., Suite 410, Washington, D.C. 20024.

FREQUENCY DISTRIBUTION—CATEGORICAL DATA	
PROJECT: *Urban Intersection 6th & D S.W.*	
PROJECT NO.: *AYZ*	
EVALUATOR/DATE: *HL 4/5/79*	
MOE: *Traffic Conflicts*	
PHASE: ☑ "BEFORE"—STEP 13 ☐ "AFTER"—STEP 18	
EVENT	FREQUENCY PER *1000 Veh.*
Left Turn	*20*
Weave	*20*
Rear-End, Amber	*24*
Rear-End, Thru-Lane	*16*
Rear-End, Left-Turn	*16*
Rear-End, Right-Turn	*14*
Rear-End, Pedestrian	*10*
Total Conflicts	*120*
Normal Maneuvers	*880*
TOTAL (Σ)	*1000*

Exhibit 3-4.
Sample Categorical Data Frequency Distribution Form

FREQUENCY DISTRIBUTION FOR "6TH AND D" TRAFFIC CONFLICTS

Conflict	Percentage of Traffic
Left Turn	2%
Weave	2%
Rear-End, Amber	2.4%
Rear-End, Through Lane	1.6%
Rear-End, Left Turn	1.6%
Rear-End, Right Turn	1.4%
Rear-End, Pedestrian	1.0%
TOTAL CONFLICTS	12%

Exhibit 3-5.
Example of a Frequency Distribution for Multi-Category Categorical Data

FREQUENCY DISTRIBUTION—ARITHMETIC SUMMARY—CLASSED DATA (CONTINUOUS)

PROJECT: *Spot Speed Study–Rural Two Lane Road*

PROJECT NO.: *EV-1*

EVALUATOR/DATE: *HL 8/17/79*

MOE: *Spot Speed*

PHASE: ☑ "BEFORE"—STEP 13 ☐ "AFTER"—STEP 18

Class			Class Frequencies (fi)	fiui	fiui²	Relative Frequencies	Cumulative Frequencies	
Boundaries	Limits	Mid-Values (ui)					Number	Relative
	28-29		0					
	30-31		1					
	32-33		2					
	34-35		14					
	36-37		7					
	38-39		20					
	40-41		38					
	42-43		29					
	44-45		35					
	46-47		15					
	48-49		12					
	50-51		9					
	52-53		4					
	54-55		0					
TOTALS (Σ)			186					

Exhibit 3-6.
Sample Continuous Data Frequency Distribution Form

Arithmetic Summaries - Arithmetic summaries are generally developed for continuous data. The two most common summary statistics are central tendency and variability. The primary central tendency summary is the average or mean value, and the main variability measure is the standard deviation. It is recommended that these two statistics be developed for all continuous MOE's. Other arithmetic summary statistics such as the median, the mode and percentiles are discussed in the section on graphic displays.

Mean - One way the mean can be calculated is from ungrouped data, as shown in Exhibit 3-7. All speed (x) values are summed (389) and dividing by N, the number of cases (14). This yields a mean of 27.8 mph.

It is often as convenient to develop the mean from grouped data (Box and Oppenlander, 1976). Exhibit 3-8 shows how this is calculated. The first thing that is done is to determine the midpoint for each class, and insert that value in the "Mid-Value" (ui) column. These values are then multiplied by the "Class Frequencies" (fi) to obtain the values in the "fiui" column. The sum, Σ"fiui" (7,877) is obtained by adding all values in the column. This figure is divided by the sum of the "Class Frequency (fi)" column (186) to obtain a mean speed of 42.3 mph.

Standard Deviation - The standard deviation, like the mean, can be calculated from unclassed data. Exhibit 3-9 shows the process. Each speed (x) is subtracted from the mean ($\overline{X}$). This quantity is squared $(\overline{X}-x)^2$ and summed. The square root of the summed deviations squared, divided by the number of measurements (N), yields the standard deviation (4.18 mph).

The method used in calculating the mean from classed data can also be used to calculate the standard deviation. Referring to Exhibit 3-10, the values in the "fiui"^{2} column are squared and inserted in the "fiui"^{2} column. This column is summed to yield 337,335. All appropriate values are inserted in the formula and the standard deviation (4.5 mph) is obtained.

ARITHMETIC SUMMARY—UNCLASSED DATA (CONTINUOUS)

PROJECT: *Caldon St.–Urban Arterial*

PROJECT NO.: *A3B*

EVALUATOR/DATE: *HL 9/4/79*

MOE: *Spot Speed*

PHASE: ☑ "BEFORE"—STEP 13 ☐ "AFTER"—STEP 18

N	x	$\bar{X} - x$	$(\bar{X} - x)^2$	N	x	$\bar{X} - x$	$(\bar{X} - x)^2$
				1	27		
				2	29		
				3	33		
				4	31		
				5	28		
				6	26		
				7	25		
				8	35		
				9	32		
				10	19		
				11	27		
				12	28		
				13	28		
				14	21		
TOTALS (Σ)				14	389		

$$\bar{X} = \frac{\Sigma x}{N} = \frac{389}{14} = 27.8 \text{ MPH}$$

$$SD = \sqrt{\frac{\Sigma (\bar{X} - x)^2}{N}} = \sqrt{\underline{\qquad}} = \sqrt{\qquad} = \underline{\qquad}$$

Exhibit 3-7.
Sample Arithmetic Summary Form
(Showing Mean Speed Calculations From Unclassed Data)

FREQUENCY DISTRIBUTION—ARITHMETIC SUMMARY—CLASSED DATA (CONTINUOUS)

PROJECT: *Spot Speed Study–Rural Two Lane Road*

PROJECT NO.: *EV-1*

EVALUATOR/DATE: *HL 8/17/90*

MOE: *Spot Speed*

PHASE: ☑ "BEFORE"—STEP 13 ☐ "AFTER"—STEP 18

Class			Class Frequencies (fi)	fiui	fiui²	Relative Frequencies	Cumulative Frequencies	
Boundaries	Limits	Mid-Values (ui)					Number	Relative
	28-29	28.5	0	0				
	30-31	30.5	1	31				
	32-33	32.5	2	65				
	34-35	34.5	14	483				
	36-37	36.5	7	256				
	38-39	38.5	20	770				
	40-41	40.5	38	1,539				
	42-43	42.5	29	1,233				
	44-45	44.5	35	1,558				
	46-47	46.5	15	689				
	48-49	48.5	12	582				
	50-51	50.5	9	455				
	52-53	52.5	4	210				
	54-55	54.5	0	0				
TOTALS (Σ)			186	7,887				

Exhibit 3-8.
Sample Arithmetic Summary Form
(Showing Mean Speed Calculations From Grouped Data)

FREQUENCY DISTRIBUTION-ARITHMETIC SUMMARY-CLASSED DATA (CONTINUOUS)

MEAN:

$$\bar{x} = \frac{\Sigma f_i u_i}{\Sigma f_i}$$

$$= \frac{7{,}877}{186}$$

$$= 42.3 \text{ MPH}$$

STANDARD DEVIATION:

$$SD = \sqrt{\frac{\Sigma f_i u_i^2 - \frac{(\Sigma f_i u_i)^2}{\Sigma f_i}}{\Sigma f_i - 1}}$$

$$= \sqrt{\frac{\quad - \frac{(\quad)^2}{\quad}}{\quad - 1}}$$

$$= \sqrt{\frac{\quad}{\quad}}$$

$$= \sqrt{\quad}$$

= __________

Exhibit 3-8.
(Continued)

ARITHMETIC SUMMARY—UNCLASSED DATA (CONTINUOUS)

PROJECT: Caldon St.–Urban Arterial

PROJECT NO.: A3B

EVALUATOR/DATE: HL 9/4/79

MOE: Spot Speed

PHASE: ☑ "BEFORE"—STEP 13 ☐ "AFTER"—STEP 18

N	x	$\bar{X} - x$	$(\bar{X} - x)^2$	N	x	$\bar{X} - x$	$(\bar{X} - x)^2$
				1	27	0.8	0.64
				2	29	1.2	1.44
				3	33	5.2	27.04
				4	31	3.2	10.24
				5	28	0.2	0.04
				6	26	1.8	3.24
				7	25	2.8	7.84
				8	35	7.2	51.84
				9	32	4.2	17.64
				10	19	8.8	77.44
				11	27	0.8	0.64
				12	28	0.2	0.04
				13	28	0.2	0.04
				14	21	6.8	46.24
TOTALS (Σ)				14	389	—	244.36

$$\bar{X} = \frac{\Sigma x}{N} = \frac{389}{14} = 27.8 \text{ MPH}$$

$$SD = \sqrt{\frac{\Sigma(\bar{X} - x)^2}{N}} = \sqrt{\frac{244.36}{14}} = \sqrt{17.45} = 4.18$$

Exhibit 3-9.
Example of Standard Deviation Calculations From Unclassed Data

FREQUENCY DISTRIBUTION—ARITHMETIC SUMMARY—CLASSED DATA (CONTINUOUS)

PROJECT: *Spot Speed Study–Rural Two Lane Road*

PROJECT NO.: *EV-1*

EVALUATOR/DATE: *HL 8/17/90*

MOE: *Spot Speed*

PHASE: ☑ "BEFORE"—STEP 13 ☐ "AFTER"—STEP 18

Class			Class Frequencies (fi)	fiui	fiui²	Relative Frequencies	Cumulative Frequencies	
Boundaries	Limits	Mid-Values (ui)					Number	Relative
	28-29	28.5	0	0	0			
	30-31	30.5	1	31	930			
	32-33	32.5	2	65	2,113			
	34-35	34.5	14	483	16,664			
	36-37	36.5	7	256	9,326			
	38-39	38.5	20	770	29,645			
	40-41	40.5	38	1,539	62,330			
	42-43	42.5	29	1,233	52,381			
	44-45	44.5	35	1,558	69,309			
	46-47	46.5	15	689	32,434			
	48-49	48.5	12	582	28,227			
	50-51	50.5	9	455	22.952			
	52-53	52.5	4	210	11,025			
	54-55	54.5	0	0	0			
TOTALS (Σ)			186	7,887	337,335			

Exhibit 3-10.
Example of Standard Deviation Calculations from Classed Data

FREQUENCY DISTRIBUTION-ARITHMETIC SUMMARY-CLASSED DATA (CONTINUOUS)
MEAN: $\bar{x} = \frac{\Sigma fi\ u_1}{\Sigma fi} = \frac{7,877}{186} = 42.3\ MPH$
STANDARD DEVIATION: $SD = \sqrt{\frac{\Sigma fi\ u_i^2 - \frac{(\Sigma fi\ u_i)^2}{\Sigma fi}}{\Sigma fi - 1}} = \sqrt{\frac{337,335 - \frac{(7,877)^2}{186}}{186 - 1}} = \sqrt{\frac{337,335 - 333,586}{185}} = \sqrt{20} = 4.5\ MPH$

Exhibit 3-10.
(Continued)

Pictorial and Graphic Displays - Data which can be summarized as a distribution or as an arithmetic summary can also be displayed pictorially or graphically. Rare event data (e.g., accidents) and categorical data (e.g., conflicts) are often displayed pictorially as Pie charts (see Exhibit 3-11) or histograms (see Exhibit 3-12). Continuous data are often displayed graphically as relative and cumulative frequency distributions.

Relative Frequency Distribution - Exhibit 3-13 uses spot speed data to show how the values needed to construct a relative frequency distribution graph are developed. The values are obtained by first dividing the number of observations for each class (the "Class Frequency (fi)") by the total number (Σ) in the sample (186). The resultant value is entered in the "Relative Frequency" column. These values are converted to percent values and plotted against the "Class Mid-Values" to produce the graph shown in Exhibit 3-15. The mode, the value with the most scores, can be obtained by inspection of the graph. In this example, it is 41 mph.

Cumulative Frequency Distribution - The procedure to develop the values to plot a cumulative frequency distribution is shown in Exhibit 3-14. The first thing that is done is to compute the "Boundaries" of the classes. The boundaries are the most extreme true value that could be included in a class and are computed to 1/2 a unit of greater precision than what would be included in a class. In this example, 1/2 mph is added to the upper limit of the class and subtracted from the lower limit of the class and inserted in the "Boundaries" column. A cumulative frequency distribution can then be developed, either from the smaller to the larger value or vice versa. When the cumulative distribution is plotted from the smaller to the larger boundary, the higher boundary is matched to the corresponding cumulative frequency distribution. When the cumulative distribution is from the larger to the smaller distribution, then the lower boundary is used. In the example, the distribution is calculated from the smaller to the larger value. In this case, the first "Cumulative Frequency-Number" and "Cumulative Frequency-Relative" are inserted in the row corresponding to the 295 boundary. The "Class Frequency (fi)" column is then summated, one line at a time in sequence, and inserted in the "Cumulative Frequency-Number" column (e.g., 1 + 2 = 3; 3 + 14 = 17, etc.). Similarly, the "Relative Frequency" column is added one line at a time in sequence and inserted into the "Cumulative Frequency-Relative" column. These values, converted to percentages, are plotted against the boundary values to develop the graph shown in Exhibit 3-16. Any percentile can be determined by inspection from this graph. For example, the first quartile (25th percentile) can be determined by inspection to be approximately 39 mph, the median (50th percentile) is 43 mph, and the 85th percentile is 47 mph.

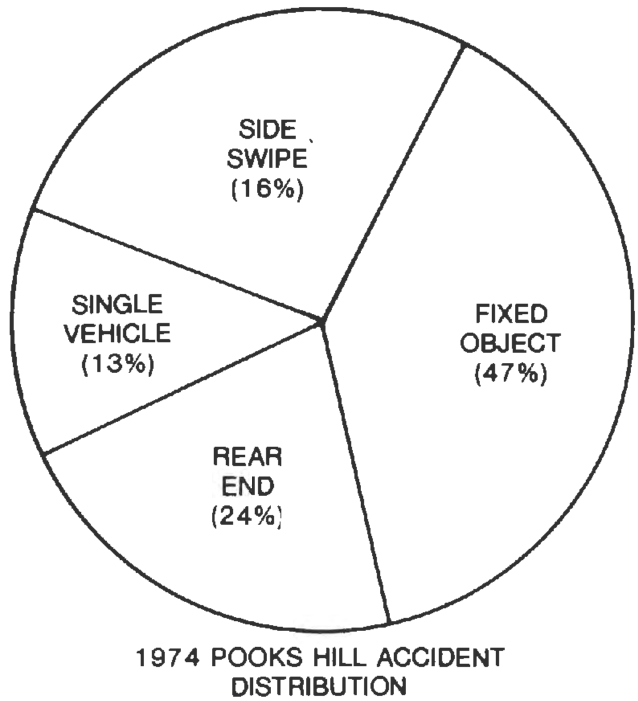


Exhibit 3-11.
Pie Chart Example

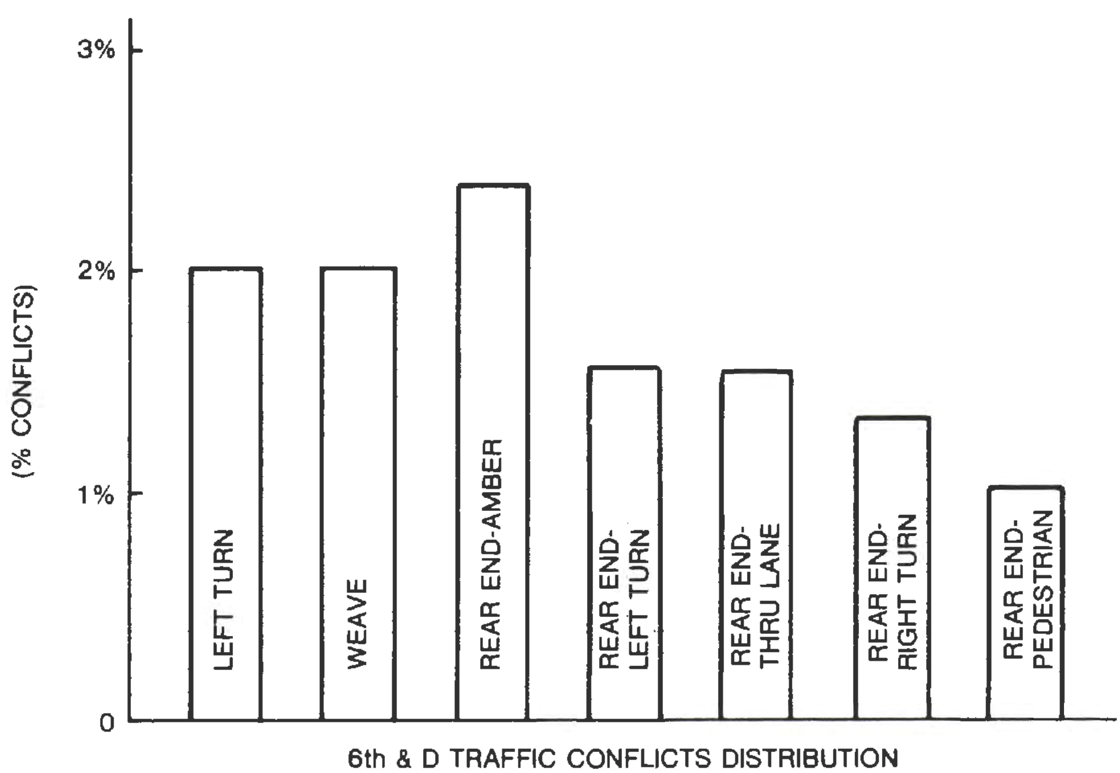


Exhibit 3-12.
Histogram Example

FREQUENCY DISTRIBUTION—ARITHMETIC SUMMARY—CLASSED DATA (CONTINUOUS)

PROJECT: *Spot Speed Study–Rural Two Lane Road*

PROJECT NO.: *EV-1*

EVALUATOR/DATE: *HL 8/17/90*

MOE: *Spot Speed*

PHASE: ☑ "BEFORE"—STEP 13 ☐ "AFTER"—STEP 18

Class: Boundaries	Class: Limits	Class: Mid-Values (ui)	Class Frequencies (fi)	fiui	$f_iu_i^2$	Relative Frequencies	Cumulative Frequencies: Number	Cumulative Frequencies: Relative
	28-29	28.5	0	0	0	0.000		
	30-31	30.5	1	31	930	0.005		
	32-33	32.5	2	65	2,113	0.011		
	34-35	34.5	14	483	16,664	0.075		
	36-37	36.5	7	256	9,326	0.038		
	38-39	38.5	20	770	29,645	0.108		
	40-41	40.5	38	1,539	62,330	0.204		
	42-43	42.5	29	1,233	52,381	0.156		
	44-45	44.5	35	1,558	69,309	0.188		
	46-47	46.5	15	689	32,434	0.081		
	48-49	48.5	12	582	28,227	0.065		
	50-51	50.5	9	455	22.952	0.048		
	52-53	52.5	4	210	11,025	0.022		
	54-55	54.5	0	0	0	0.000		
TOTALS (Σ)			186	7,887	337,335	1.000		

Exhibit 3-13.
Example of Calculations for Relative Frequency Distributions

FREQUENCY DISTRIBUTION—ARITHMETIC SUMMARY—CLASSED DATA (CONTINUOUS)

PROJECT: *Spot Speed Study–Rural Two Lane Road*

PROJECT NO.: *EV-1*

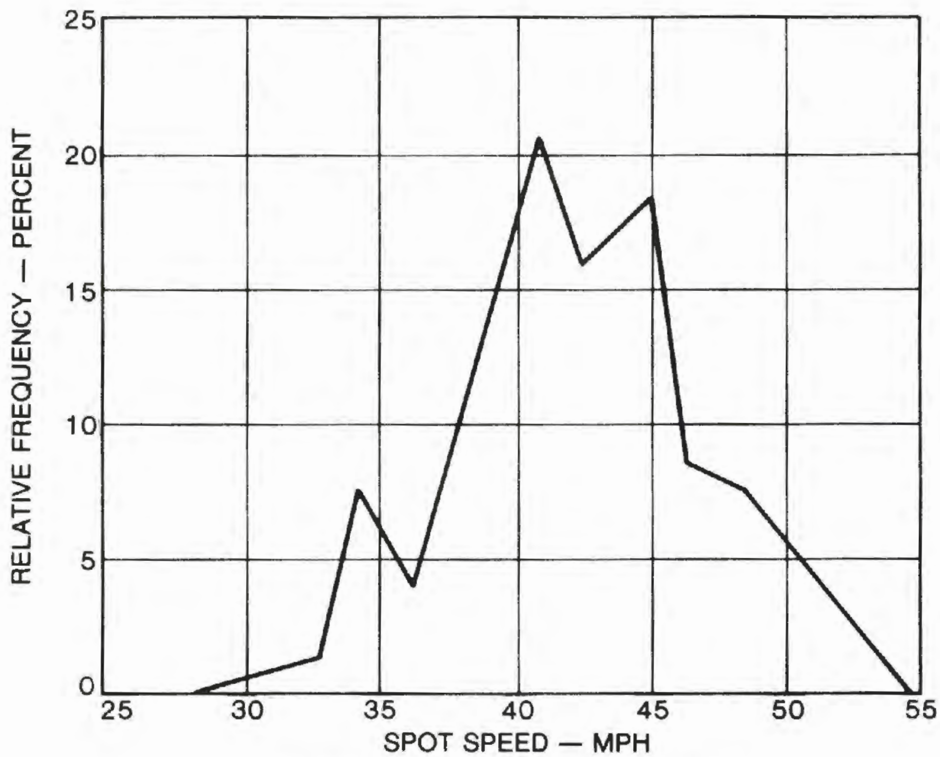
EVALUATOR/DATE: *HL 8/17/90*

MOE: *Spot Speed*

PHASE: ☑ "BEFORE"—STEP 13 ☐ "AFTER"—STEP 18

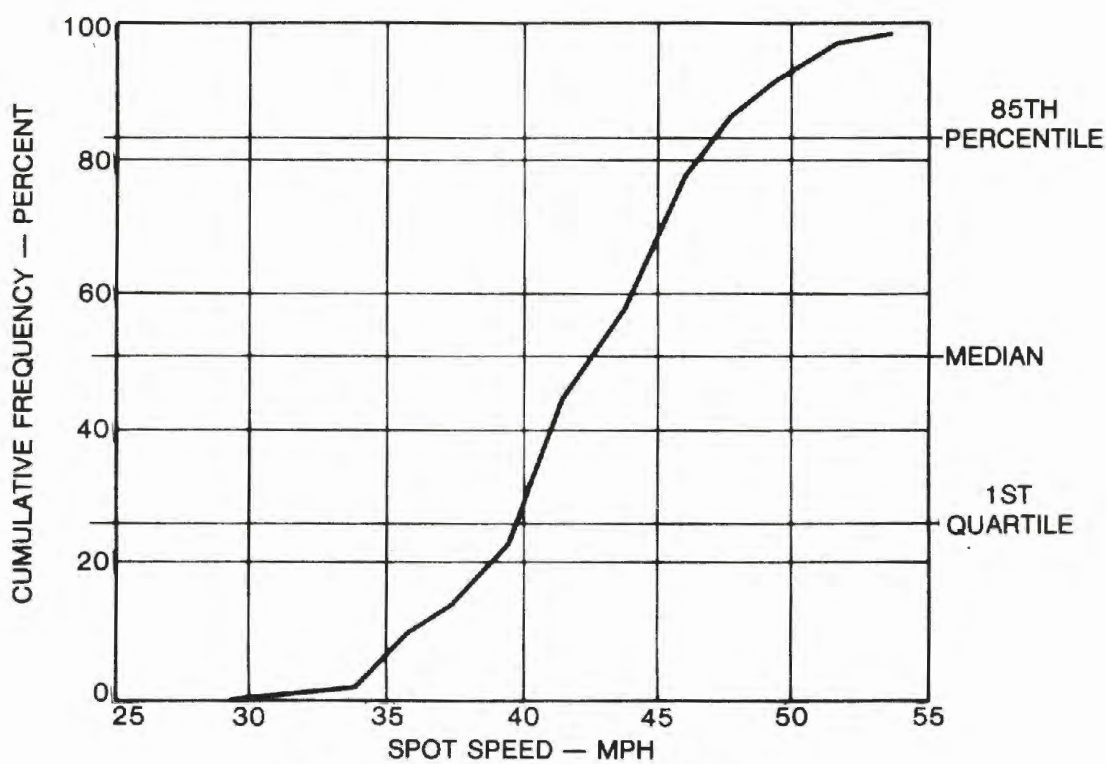
Class: Boundaries	Class: Limits	Class: Mid-Values (ui)	Class Frequencies (fi)	fiui	fiui²	Relative Frequencies	Cumulative Frequencies: Number	Cumulative Frequencies: Relative
27.5							0	0.000
	28-29	28.5	0	0	0	0.000		
29.5							1	0.005
	30-31	30.5	1	31	930	0.005		
31.5							3	0.016
	32-33	32.5	2	65	2,113	0.011		
33.5							17	0.092
	34-35	34.5	14	483	16,664	0.075		
35.5							24	0.129
	36-37	36.5	7	256	9,326	0.038		
37.5							44	0.237
	38-39	38.5	20	770	29,645	0.108		
39.5							82	0.441
	40-41	40.5	38	1,539	62,330	0.204		
41.5							111	0.597
	42-43	42.5	29	1,233	52,381	0.156		
43.5							146	0.790
	44-45	44.5	35	1,558	69,309	0.188		
45.5							161	0.866
	46-47	46.5	15	689	32,434	0.081		
47.5							173	0.936
	48-49	48.5	12	582	28,227	0.065		
49.5							182	0.979
	50-51	50.5	9	455	22.952	0.048		
51.5							186	1.000
	52-53	52.5	4	210	11,025	0.022		
53.5								
	54-55	54.5	0	0	0	0.000		
55.5								
TOTALS (Σ)			186	7,887	337,335	1.000	186	1.000

Exhibit 3-14.
Example of Calculations for Cumulative Frequency Distributions



RURAL SPOT SPEED DISTRIBUTION

Exhibit 3-15.
Relative Frequency Distribution Example



RURAL SPOT SPEED DISTRIBUTION

Exhibit 3-16.
Cumulative Frequency Distribution Example

STEP 14. Schedule Improvement

Steps 14 and 15 taken together form the implementation phase. This phase begins after assurances have been made that the "Before" data are acceptable. Prior to starting this phase, it would be useful to look over what has been accomplished and to be assured that the project has all information and materials for improvement implemention.

The first step consists of two activities, scheduling and implementing the improvement. While there is no hard and fast rule for when the improvement should be implemented, it is recommended that the application begin as soon as possible after the "Before" data are collected and the improvement developed to reduce the possibility of changes over time. Figure 3-5 shows the Step 14 flow.

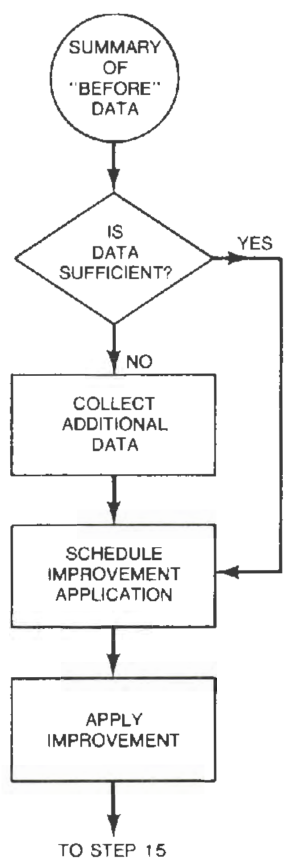


Figure 3-5.
Flow. Step 14 - Schedule Improvement

STEP 15. Allow for Acclimation

The Step 15 task depends, to a large extent, on the implementation and acclimation time period, between "Before" and "After" data collection phases. If the acclimation portion is relatively short, then all that generally is required is to allow the time to pass. If the total time period (including implementation) is long (greater than 1 year), or if changes over time may occur, then site monitoring will be required and applicable controls, specified in Step 5, should be applied.

When site monitoring is deemed necessary, such things as volume changes, new construction, or traffic pattern shifts should be watched. If changes are occurring, or about to occur, their effects should be gauged and a determination made as to whether the acclimation portion should be modified, shortened or terminated. The strategies in Step 5 (Table 2-8) will serve as a guide. In any event, all changes must be documented.

Figure 3-6 shows the Step 15 flow. A sample Acclimation Assessment form is shown in Exhibit 3-17.

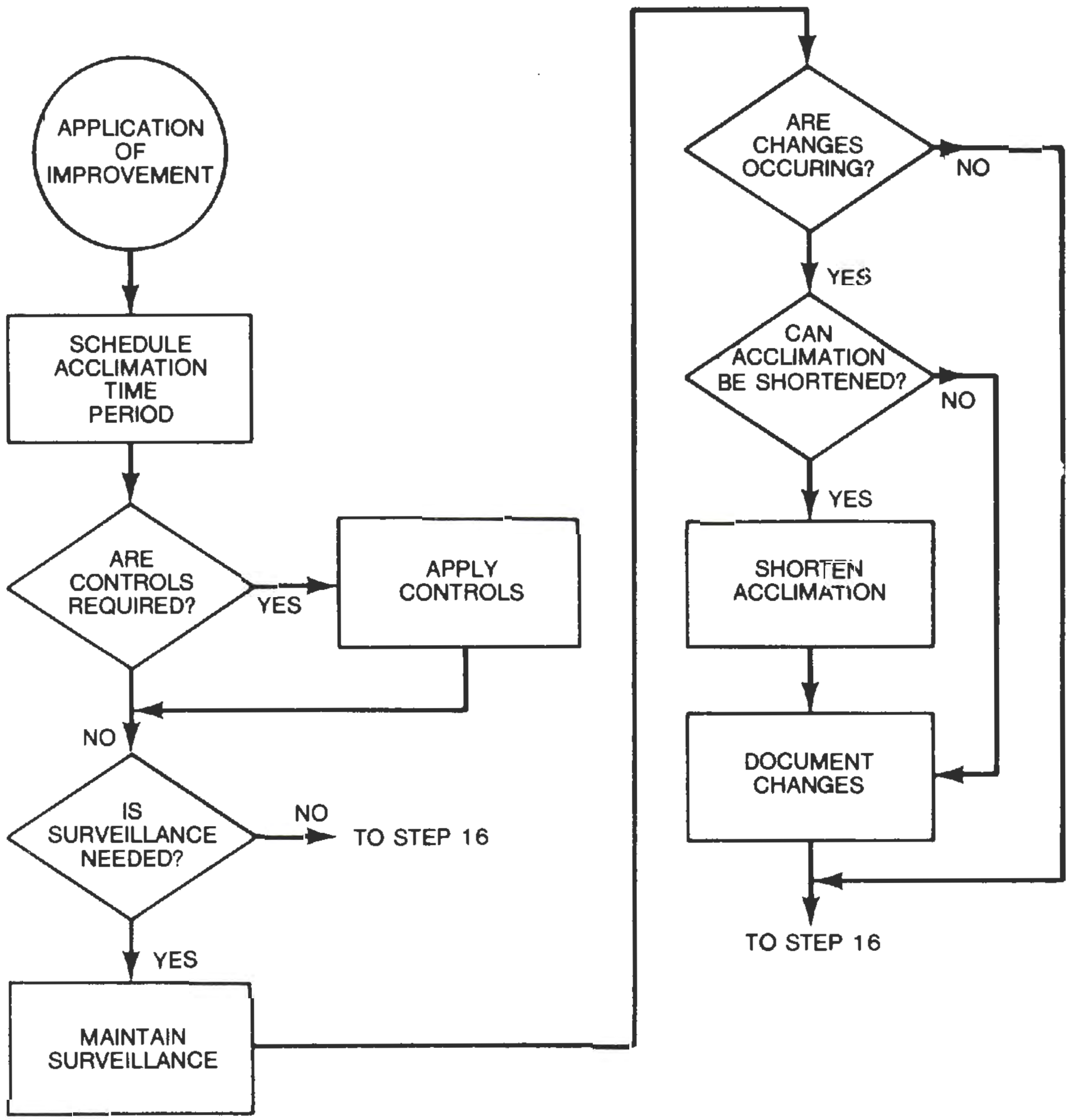


Figure 3-6.
Flow, Step 15 - Allow for Acclimation

ACCLIMATION ASSESSMENT—STEP 15

PROJECT: *Interchange of 7th Street Bypass*

PROJECT NO.: *1234B*

EVALUATOR/DATE: *HL 4/15/80*

IMPLEMENTATION LENGTH: *4 months*

ACCLIMATION LENGTH: *8 months*

TOTAL LENGTH: *1 year*

POSSIBLE CHANGES OVER TIME	CONTROLS
– *Enforcement*	*Coordinate with State H.P.*
– *Main Street Mall*	*Coordinate with Planning*
– *Volume change next summer*	*Shorter Acclimation to less than 1 year*

SURVEILLANCE REQUIREMENTS:

- *Monitor Volume*
- *Monitor Enforcement*

SURVEILLANCE RESULTS:

- *Large increase in volume noted & expected*
- *Police to more closely enforce 55 MPH*

REQUIRED CHANGES:

- *Shorten Acclimation to end April 1980, Schedule "Pre-Data Collection" for 4/17/80*

Exhibit 3-17.
Sample Acclimation Portion Assessment Form

STEP 16. Perform Pre-Data Collection ("After" Phase)

The basic assumption upon which the evaluation is based (given a Before/After design) is that the only thing about the site to change is the improvement. Therefore, the "After" phase must be comparable to the "Before" phase. As a matter of fact, the most important consideration in the post-implementation phase, irrespective of which evaluation design is used, is to assure comparability with "Before" conditions. With this in mind, the primary task in Step 16 is to perform those pre-data collection tasks required to assure that "After" data are collected in a manner as close to "Before" data collection as possible.

The first Pre-Data Collection ("After" phase) activity is to assess all aspects of the evaluation to determine whether and what changes to the Evaluation Plan are required for the "After" phase. Hopefully, nothing will have transpired in either the "Before" data collection step or the implementation and acclimation period that would necessitate changing the "After" data collection plan. However, it is possible that changes in conditions, equipment, personnel or site characteristics will have occurred. If so, these changes must be reflected in the "After" Data Collection step to maintain comparability.

Even when no major changes have occurred, there usually are a number of routine procedures or minor changes requiring some pre-data collection activities. For example, new or different equipment may be used, requiring scheduling and procurement. In addition, recalibration may have to be scheduled or additional comparability checks made. In a similar manner, new data collection personnel will require training. Retraining of original project personnel may also be required in the normal course of events. The same is true for forms, coordination, or procedures. Any of these factors, be they due to major changes, minor changes, or routine procedures, may also necessitate pilot testing. If such is the case, details of the pilot tests must be worked out and the tests conducted prior to collecting "After" data.

The Step 16 flow is shown in Figure 3-7. Prior to Step 17, personnel should fill out a Pre-Data Collection form such as the one shown in Exhibit 3-1. They should also review all logs, the "Before" data collection results and the acclimation surveillance results in order to make all necessary modifications.

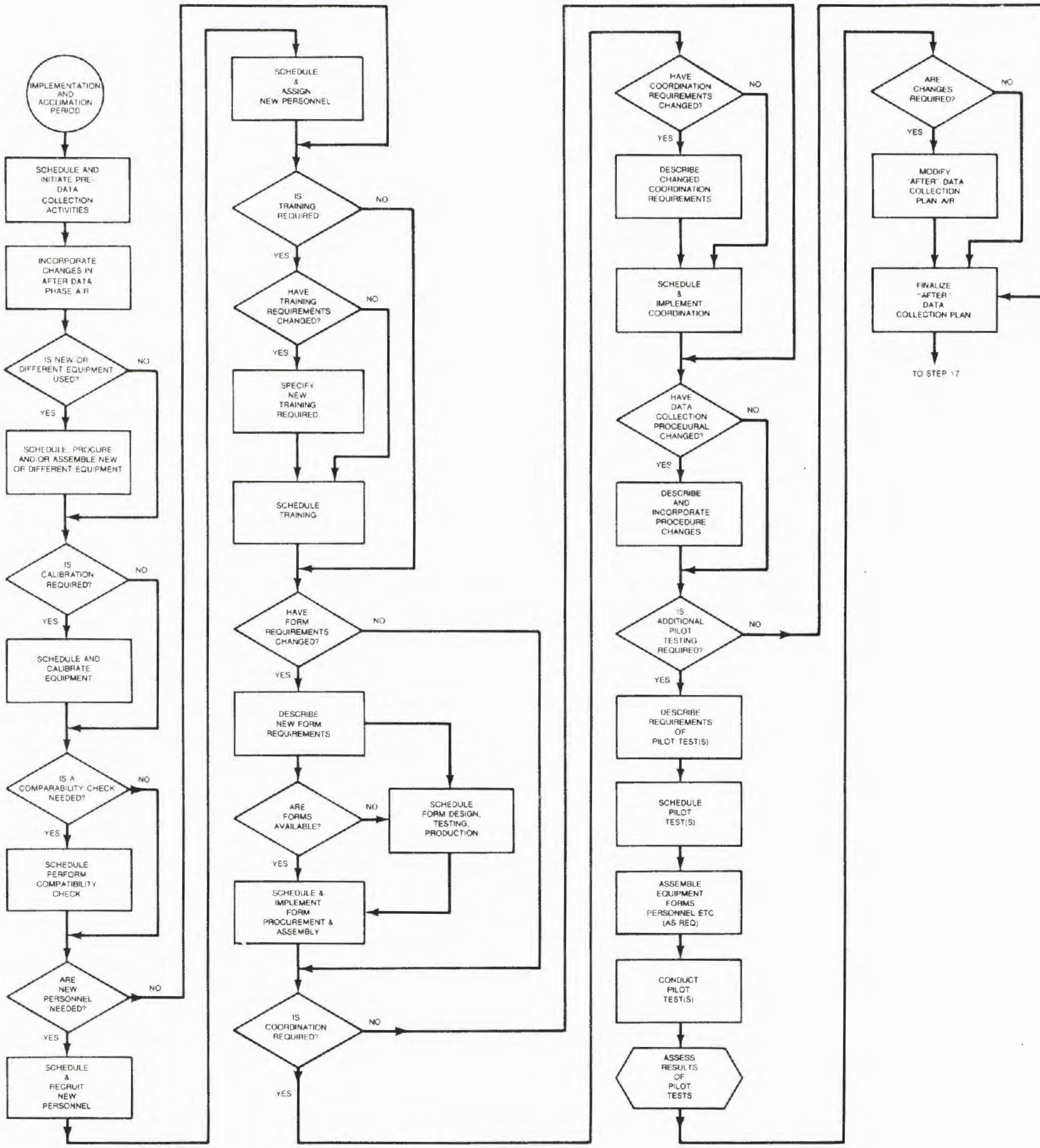


Figure 3-7.
Flow, Step 16 - Perform Pre-Data Collection (''After Phase)

STEP 17. Collect "After" Data

The Step 17 activities are essentially identical to those of Step 12. That is, equipment, personnel, forms, etc., are assembled, data are collected using appropriate sampling techniques, and data quantity and quality are assessed. The Step 17 flow is presented in Figure 3-8.

An additional task in the "After" data collection step is to assure comparability. Basic environmental and traffic stream characteristics should be assessed, and a determination made that "Before" conditions are essentially replicated. If deviations are found, project personnel should attempt to modify the "After" data collection step, e.g., by delaying data collection or adjusting other aspects. If comparability still cannot be achieved, then unavoidable instances of noncomparability must be documented and their effect on the overall evaluation determined.

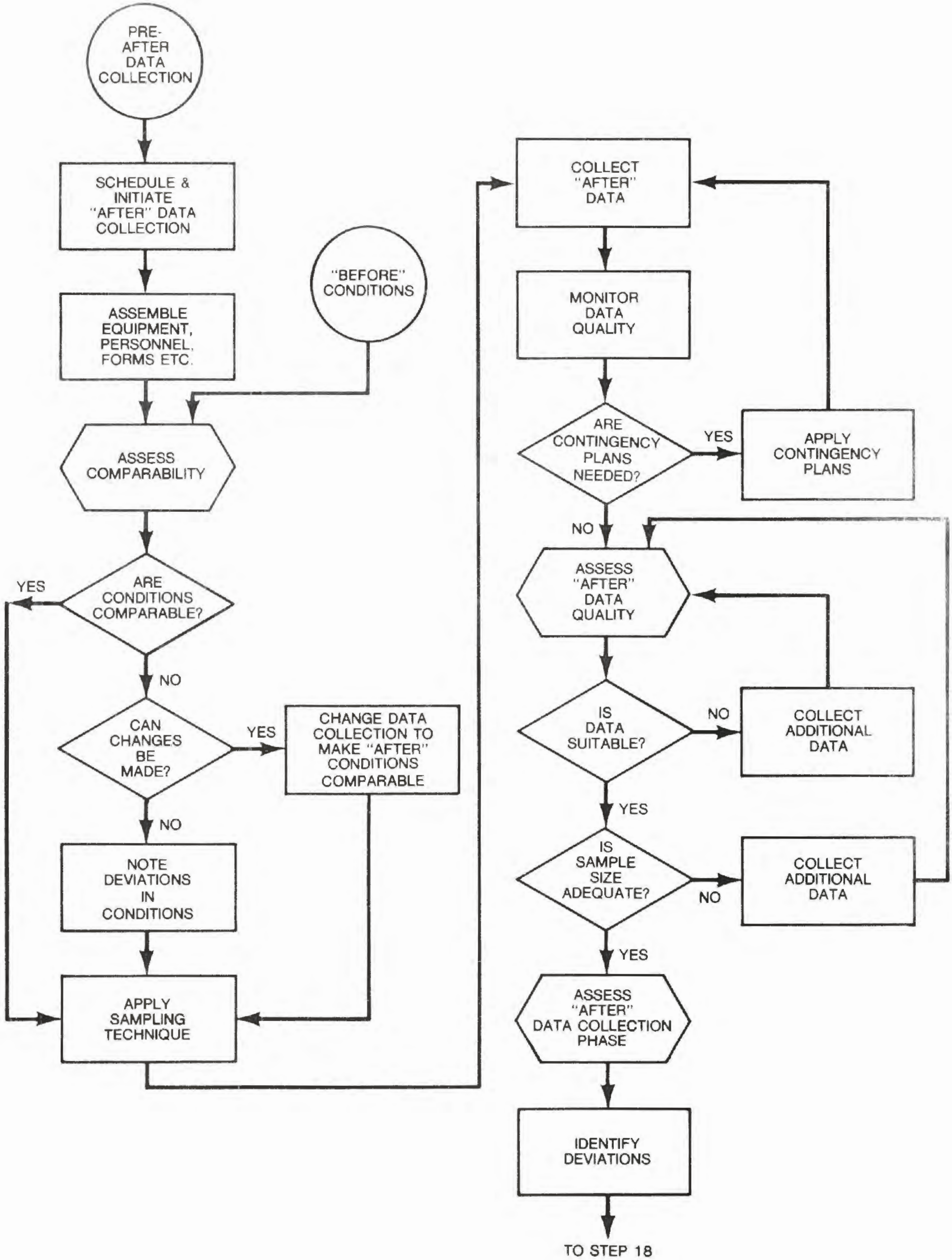


Figure 3-8.
Flow, Step 17 - Collect "After" Data

STEP 18. Reduce and Summarize "After" Data

The final step in the post-implementation phase, shown in Figure 3-9, is to reduce and summarize the "After" data. This step is the same as Step 13. Personnel should use the same forms and develop the same descriptive statistics, so that comparisons can be made in the Results Assessment phase.

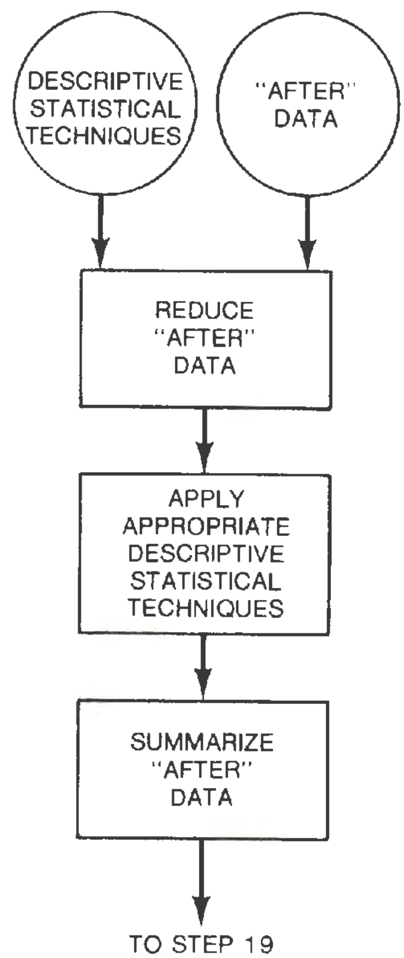


Figure 3-9.
Flow, Step 18 - Reduce and Summarize "After" Data

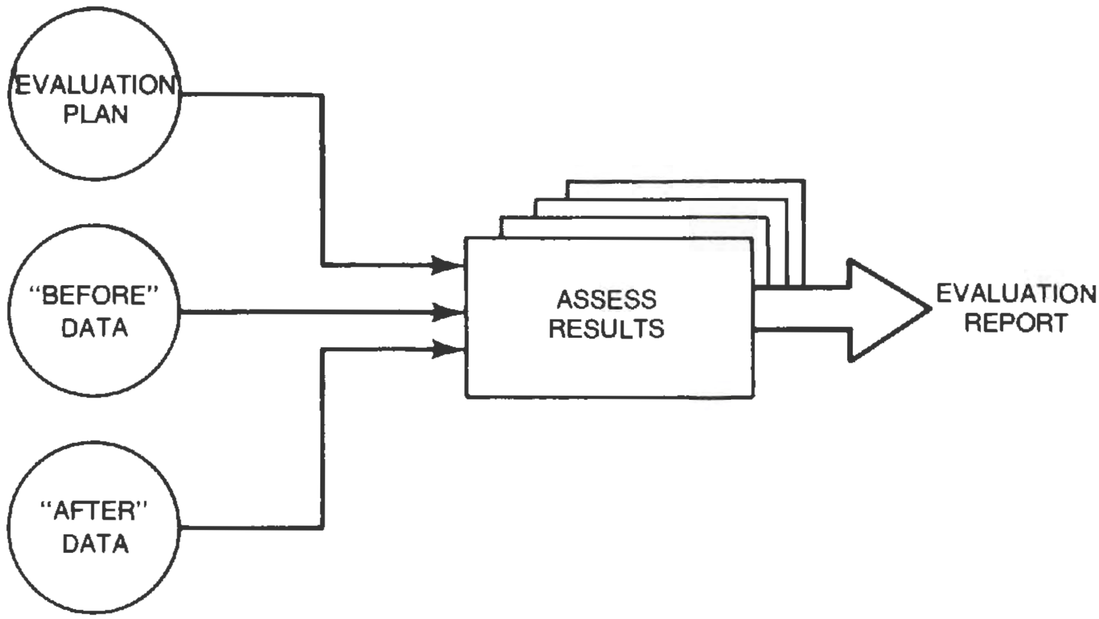
ASSESSMENT OF RESULTS

This phase of the evaluation process marks the culmination of the effort, where all aspects of the procedure are gathered together and an assessment made of the results. While the effectiveness determination rests largely on whether the differences in the data are statistically significant, there are other factors that are considered. One of these is whether the results are practically significant. Another is whether the evaluation was conducted properly. The last is to determine whether conditions or events beyond the control of the evaluator may have led to spurious effects.

As in the planning of the evaluation, it is recommended that users unfamiliar with statistics obtain the services of a statistician to aid in performing the statistical tests and interpreting the results. It may also be useful to enlist the assistance of an individual who is knowledgeable in highway engineering economy when an economic evaluation is appropriate and when such knowledge is not possessed by the evaluator.

Overview

Figure 4-1 presents a functional flow of the steps used to assess the results of the evaluation. All phases of the evaluation process input into this task. The Evaluation Plan provides information relative to what tests of significance to use, and what level of significance to test to. The Collection of Evaluation Data provides the "Before" and "After" data used to make the comparisons to test for significance. If the results are found to be statistically significant, then practical significance is evaluated. The final step involves forming conclusions and recommendations and reporting the findings. Continuity of steps is continued in this phase, with the first step beginning at Step 19.



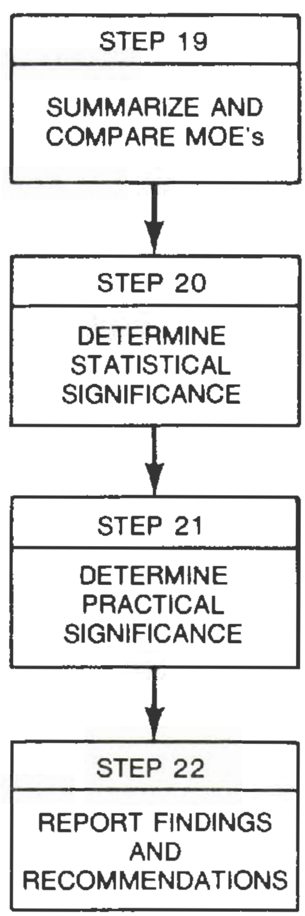


Figure 4-1.
Result Assessment - Functional Flow

STEP 19. Summarize and Compare Traffic and System Performance MOE's

Prior to applying applicable tests of significance, a necessary first step is to summarize "Before" and "After" data, tabulate it for comparison, and determine differences in the MOE's. This information provides an indication of the magnitude of the difference, and thereby aides in assessing the practical importance of the improvement. The comparison is also used to formulate conclusions and recommendations.

Figure 4-2 shows the Step 19 flow. Data used for MOE comparisons comes from Step 13 (Reduce and Summarize "Before" Data) and Step 18 (Reduce and Summarize "After" Data). This step discusses Traffic and System Performance MOE's. Accident summaries and comparisons are discussed in Step 19A.

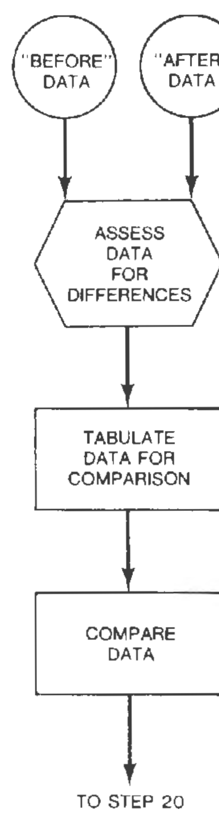


Figure 4-2.
Flow. Step 19 - Summarize and Compare MOE's

In developing summaries and comparisons for Traffic and System Performance MOE's, the assumption is made that since the time period between the "Before" and "After" data collection period is short, in most instances changes in volume will not affect the MOE. In addition, because most categoried MOE's are rate related, minor changes in volume will be washed out by "Before" and "After" MOE's being expressed in the same volume units. For continuous MOE's such as speed, small volume changes should not have enough of an effect on the mean to make a difference, although volume changes may affect variance. Finally, since major volume changes will adversely affect comparability, such changes should have been noted and taken into account in Steps 16 and 17.

In Step 19, the descriptive statistics developed in Steps 13 and 18 are used to summarize "Before" and "After" data and to make comparisons of the differences. Generally, both the magnitude of the differences and the percentage changes are useful statistics to develop.

Exhibit 4-1 provides a form to use with categorical data. A formula for calculating the percent change for rate-related MOE's is presented on this form; where (E_R) = Expected rate at the project or control site if the improvement was not made, and (A_R) = "After" MOE rate. Since the assumption is made that the "Before" MOE rate would not have changed if no improvement were made, (E_R) equals the "Before" project or control site rate. In the example shown, the "Before" project erratic maneuver rate is 47 erratic maneuvers per 1,000 vehicles, and the "After" project erratic maneuver rate is 28 erratic maneuvers per 1,000 vehicles. This yields a 19 erratic maneuver per 1,000 vehicle reduction. Substituting the data in the formula, this reduction represents a 43% change.

Other comparisons which can be made are in the form of contingency tables comparing both categorical (Exhibit 4-2) and continuous (Exhibit 4-3) MOE's. Graphic comparisons are also often made for categorical (Exhibit 4-4) and continuous (Exhibit 4-5) MOE's.

MOE COMPARISON—CATEGORICAL DATA—STEP 19				
PROJECT: *Diagrammatic Guide Signs–I 70S*				
PROJECT NO.: *I 70S-1*				
EVALUATOR/DATE: *HL 3/25/80*				
DESIGN: ☑ BEFORE-AFTER ☐ BEFORE-AFTER WITH A CONTROL SITE				
MOE: *Erratic Maneuver*	Project		Control Site	
Event	Before	After	Before	After
	(B_{PR})	(A_{PR})	(B_{CR})	(A_{CR})
Exiting Traffic	*47*	*28*		

NOTE:
FOR BEFORE/AFTER

$E_R = B_{PR}$
$A_R = A_{PR}$

FOR CONTROL SITE
$E_R = B_{PR}\ (A_{CR}/B_{CR})$
$A_R = A_{PR}$

$$\% \text{ CHANGE} = \left[\frac{E_R - A_R}{E_R}\right] \times 100$$

$$= \left[\frac{47-28}{28}\right] \times 100$$

$$= 43\%$$

Exhibit 4-1.
Sample Categorical Data MOE Comparison Form

DIAGRAMMATIC GUIDE SIGN IMPROVEMENT
(Erratic Maneuvers per 1,000 Vehicles)

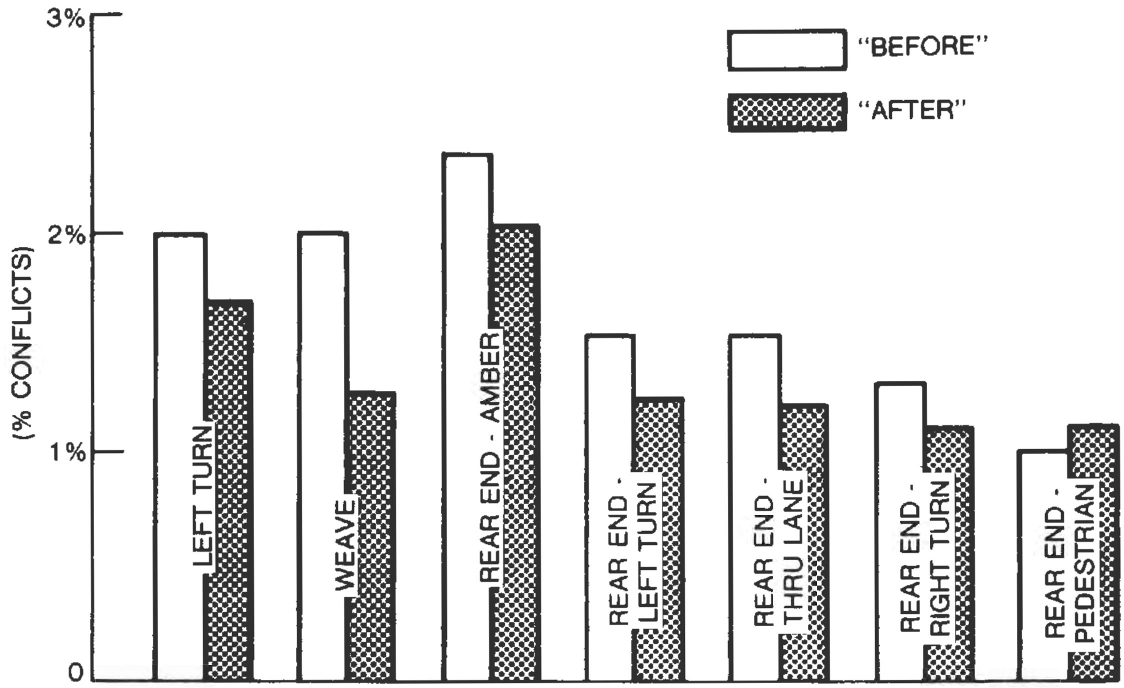
EXITING TRAFFIC		THRU TRAFFIC	
"Before"	"After"	"Before"	"After"
47	28	72	39

Exhibit 4-2. Categorical MOE Frequency Distribution Comparison Example

FREEWAY SPOT SPEED STUDY

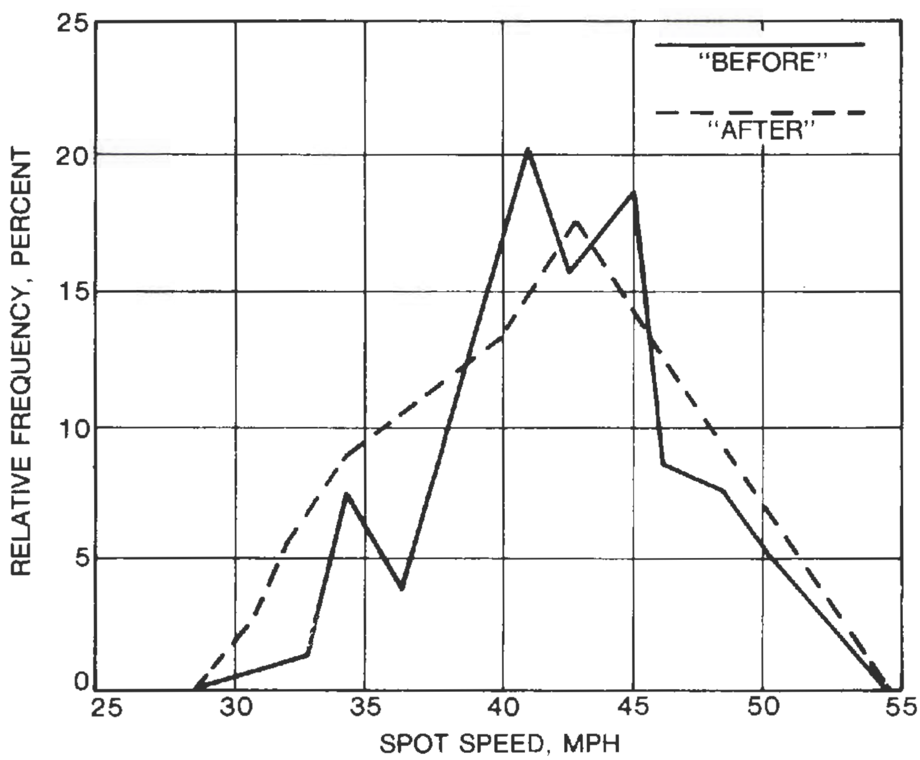
(MPH)					
15 Percentile		Median		85 Percentile	
"Before"	"After"	"Before"	"After"	"Before"	"After"
38	35	59	55	63	60

Exhibit 4-3. Continuous MOE Frequency Distribution Comparison Example



6th & D TRAFFIC CONFLICTS DISTRIBUTION, BEFORE AND AFTER A SIGNAL TIMING PROJECT

Exhibit 4-4.
Categorical MOE Graphic Comparison Example



RURAL SPOT SPEED DISTRIBUTIONS, BEFORE AND AFTER SPEED SIGNING

Exhibit 4-5.
Continuous MOE Graphic Comparison Example

Step 19A. Summarize and Compare Accident MOE's.

Appendix B presents the procedure to use when summarizing and comparing accident MOE's. Because there is a long data accumulation period following the "Before" phase, and because accident MOE's are sensitive to changes in traffic volume, it will be necessary to estimate the expected MOE values prior to testing for significance. The expected value is derived differently for each plan (evaluation design). It also depends on the nature of the data (i.e., whether it is frequency-related or rate-related). Finally, it depends on the characteristics of the MOE over time.

The procedure to follow is to first prepare a summary table for all data, then to calculate the percent change in the MOE by selecting the appropriate plan and applying the appropriate formulas for either frequency or rate-related MOE's. The expected values of the MOE's are estimated using a linear regression technique if the yearly values of the MOE follow an increasing or decreasing trend. If the MOE values follow a horizontal trend, or are widely dispersed, the mean value of the MOE over the entire analysis period is used for the expected MOE estimation. The reader should refer to Datta et al. (1978) for a full description of the procedure. Appendix B, taken from that report, contains the needed formulas.

The following examples have been taken from Datta et al. (1978).

Comparison of MOE's Using the Before-After with Control Site - A highway safety project site consists of a two-lane, 2-mile long highly traveled roadway section with a number of sharp curves. A majority of accidents at this site during the "Before" period were of the "run-off-the-road" type. The safety project implemented at this site included straightening these curves through major reconstruction. The entire section was then edgelined.

A single control site was identified by the evaluator prior to project implemention and the Before-After with Control Site plan was used. The objectives of the evaluation were to determine the effect of the project on total accidents, total run-off-the-road (ROR) accidents, total fatal and total personal injury accidents. The MOE's are rate-related and reflect the percent change in each of the objectives. The data summary table for this project is shown in Exhibit 4-6.

ACCIDENT DATA COMPARISON—STEP 19 A

PROJECT: *Two-Lane, Two-Mile Long Roadway*

PROJECT NO.: *C-1*

EVALUATOR/DATE: *8/22/77 MUL*

PLAN: ☐ BEFORE/AFTER ☑ CONTROL SITE

Data Summary	Control Before (B_{CF})	Control After (A_{CF})	Project Before (B_{PF})	Project After (A_{PF})	Expected After Rate *X* or Freq. ___	Percent Reduction (%)
Accidents:						
(Fundamental)						
(3 years)						
Total Accidents	*30*	*21*	*24*	*21*		
Fatal Accidents	*9*	*6*	*12*	*3*		
Injury Accidents	*12*	*6*	*12*	*6*		
PDO Accidents	*9*	*9*	*0*	*9*		
(Project Objectives)						
Total ROR Accidents	*15*	*12*	*12*	*9*		
Exposure *(3 years)*						
units: ___ V, or *M* VM	*5.01*	*5.37*	*3.93*	*4.74*		
Comparison Rate *X* **or Frequency** ___	B_C *R*	A_C *R*	B_P *R*	A_P *R*	E *R*	(%)
Total Accidents/ *MVM*	*5.99*	*3.91*	*6.11*	*4.43*	*3.99*	*−11.0*
Fatal Accidents/ *MVM*	*1.80*	*1.12*	*3.05*	*0.63*	*1.90*	*66.8*
Injury Accidents/ *MVM*	*2.40*	*1.12*	*3.05*	*1.27*	*1.42*	*10.6*
PDO Accidents/ *MVM*	*1.80*	*1.68*	*0*	*2.53*	*0*	—
Total ROR/MVM	*2.99*	*2.23*	*3.05*	*1.90*	*2.27*	*16.3*

Exhibit 4-6.
Sample Accident Data Comparison Form (Control Site Example)

The data were plotted, and it was determined that there was no trend in the accidents over time. Therefore, the regression analysis technique was not selected.

For illustrative purposes, only total ROR accidents are considered. The expected value and the percentage change in the total ROR accident rate are calculated below:

$E_R = B_{PR} (A_{CR}/B_{CR})$

Where:

E_R = Expected rate-related MOE

B_{PR} = "Before" period MOE rate at the project site

A_{CR} = "After" period MOE rate at the control site

B_{CR} = "Before" period MOE rate at the control site

Substituting into the above equation:

$E_R = 3.05 (2.23/2.99)$

= 2.27 ROR accidents/million vehicle-miles

and

A_{PR} = "After" period total ROR accident rate at the project site

= 1.90 accidents/million vehicle-miles.

Percent Change = $((E_R - A_{PR})/E_R)100$

= ((2.27 - 1.90)/2.27)100

= 16.3% decrease in accidents/million vechicle-miles of travel.

Expected "Before" accident frequencies were calculated using the following equation:

$$E_F = E_R \times (\text{"After" Project Exposure})/10^6$$

Where:

E_F = Expected "Before" accident frequency

E_R = Expected rate-related MOE if the improvement had not been made.

The expected 3-year "Before" accident frequency for ROR accidents was calculated using the appropriate values from the above sample calculations and Exhibit 4-6.

$$E_F = 2.27 \times 4.74 \text{ MVM}$$
$$= 10.8 \text{ ROR accidents for 3 years.}$$

Similar calculations should be performed for the remaining data contained in Exhibit 4-6 to determine the percent change for each data variable entry and the expected "Before" accident frequency.

Comparision of MOE's Using the Before-After - On the northbound approach of a high volume signalized intersection near a steep downgrade, a large number of rear-end accidents were observed during the "Before" period. The safety project implemented at this site consisted of increasing the amber time by two seconds and installing advance warning signs. It was not possible to identify control sites for this project and the Before-After plan was selected. The purpose of the project was to reduce rear-end accidents and severity of accidents at the intersection. The evaluation objectives were to determine the effect of the project on total rear-end, total fatals and total personal injury accidents. The data summary table for 3 years of data for this example is shown in Exhibit 4-7.

ACCIDENT DATA COMPARISON—STEP 19 A

PROJECT: *Urban Signalized Intersection*

PROJECT NO.: *C-2*

EVALUATOR/DATE: *8/22/77 RUR*

PLAN: ☑ BEFORE/AFTER ☐ CONTROL SITE

Data Summary	~~Control~~ ~~Before~~ (B_{CF})	~~After~~ (A_{CF})	Project Before (B_{PF})	Project After (A_{PF})	Expected After Rate *X* or Freq. ___	Percent Reduction (%)
Accidents: *(3 years)*						
(Fundamental)						
Total Accidents			*55*	*21*		
Fatal Accidents			*4*	*1*		
Injury Accidents			*38*	*14*		
PDO Accidents			—	—		
(Project Objectives)						
Rear-End Accidents			*40*	*10*		
Exposure *(3 years)*						
units: *M* V, or ___ VM			*11.00*	*11.50*		
Comparison Rate ___ or Frequency ___	~~B_C ___~~	~~A_C ___~~	B_P *R*	A_P *R*	E *R*	(%)
Total Accidents/ *MV*			*5.50*	*1.83*	*5.50*	*66.7*
Fatal Accidents/ *MV*			*0.40*	*0.09*	*0.40*	*77.5*
Injury Accidents/ *MV*			*3.80*	*1.22*	*3.80*	*67.9*
PDO Accidents/ *MV*			—	—	—	—
Rear-End Accidents/MV			*4.00*	*0.87*	*4.00*	*78.2*

Exhibit 4-7.
Example of Before/After Accident Summary and Data Comparison

For illustrative purposes, only personal injury accidents are considered. The expected value and the percent change in the personal injury accident rate, using the injury accident rates shown in Exhibit 4-7, is shown below:

$E_R = B_{PR} = 3.80$ personal injury accidents/MV

Where:

E_R = Expected rate-related MOE if the improvement had not been made

B_{PR} = "Before" period MOE rate at the project site

and

Percent Change $= ({}^{E}R - {}^{A}PR)/{}^{E}R \ \ 100$

Where:

A_{PR} = "After" period MOE rate at the project site = 1.22 personal injury accidents/MV

= (3.80 - 1.22)/(3.80)(100)

= 68% decrease in personal injury accidents/MV.

Expected "Before" accident frequencies were calculated as follows:

$E_F = E_R$ x "After" Project Exposure/10^6

Where:

E_F = Expected "Before" accident frequency

${}^{E}R$ = Expected rate-related MOE if the improvement had not been made.

Substituting for the injury accident MOE, the expected 3-year "Before" injury accident frequency was calculated as follows:

E_F = 3.8 x 11.5 MV

= 43.7 injury accidents for 3 years.

The percent changes and expected "Before" frequencies were similarly calculated for the remaining accident types.

STEP 20 <u>Determine Statistical Significance</u>

The primary determination of the effectiveness of the project is based on whether or not the results of the evaluation achieve statistical significance. Statistical significance is established by testing the differences between the "Before" and "After" data and determining the probability of the results being due to chance. If it can be established that the confidence in the results not being due to chance is equal to or greater than the established acceptance value, then the results are deemed significant. Once the tests to be used have been established and the confidence level to be tested to determined, then the actual testing is simply a matter of inserting values into a figure or formula and reading results from the figure or a significance table. Since the rationale for selecting a test and confidence level has been discussed in the Evaluation Plan section, this section will focus on how to apply the various tests.

While there are a number of statistical tests which can be used to test for significance, five basic tests are recommended to assess most evaluation results. These tests, Poisson, t, F, Z, and Chi Square, are presented in Steps 20A through 20E. In addition, while these tests can be applied at any confidence level, 80% (.20), 90% (.10) and 95% (.05) are the most likely to be used. Accordingly, all examples are limited to these five tests and three levels. Figure 4-3 shows the Step 20 flow.

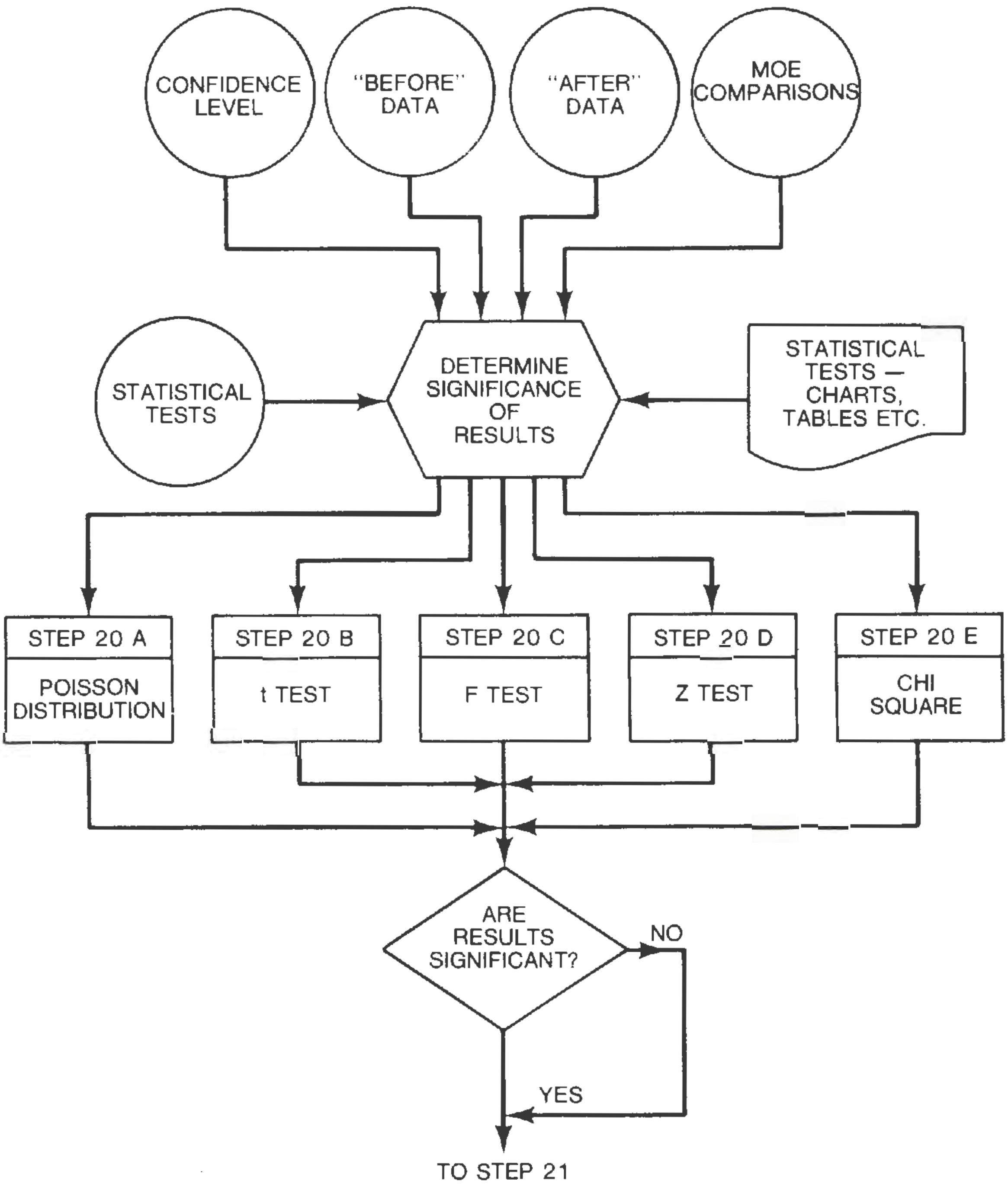


Figure 4-3.
Flow, Step 20 - Determine Statistical Significance

STEP 20A. Poisson Distribution Test

The recommended test to use to determine whether reductions in accident MOE's are significant is the Poisson Distribution test. This test is based on the premise that accidents are rare events that are distributed in accordance with the Poisson Distribution (a special case of the binomial distribution where the probability of any event is extremely small).

Because accidents assume a Poisson Distribution, differences between the average number of accidents of two samples, for example "Before" accidents and "After" accidents, randomly selected from a common distribution, will have known characteristics which can be defined by a given Poisson Distribution curve. If it can be concluded that the "Before" and "After" results are from different distributions, then it can be assumed that the improvement had an effect on the accident experience at the site. When it is found that the differences are from the same distribution, it is assumed that the improvement had no effect. This determination is made by entering the differences on the curve for the selected confidence level and seeing whether or not the results fall above (i.e., a different distribution and therefore significant) or below (i.e., the same distribution and therefore not significant) the curve.

A 3-year "Before" and 3-year "After" data collection period is recommended, since it is necessary to accumulate data over a multi-year period to use the Poisson test. This also achieves a larger data base which might otherwise not be forthcoming, particularly at sites with low traffic volumes. The nature of the Poisson curves is such that the percentage reduction in accidents needed to obtain significant increases as accidents at a site decrease. Thus, if the number of "Before" accidents at a site is low to begin with, then the percent change in accidents in the "After" phase will have to be high.

Exhibit 4-8 provides an example, taken from Datta et al (1978), illustrating how the Poisson Distribution test is applied. The example involved the installation of guardrail to shield roadside obstacles. A high confidence level of 95% was felt to be justified.

The expected "Before" accident frequencies for the 3 years and the percent change in each MOE category were first tabulated in the appropriate columns in the form shown in Exhibit 4-8. To illustrate how the required percent reduction column is developed, the example was worked out for "total accidents." The Poisson curve shown on page 2 of the exhibit was entered on the horizontal axis with the expected 3-year "Before" total accident frequency (176.6). Reading the vertical axis at its intersection with the 95% curve yields the required percent reduction of 12%. Since the observed change of -1.4% was less than the required change, the change was not significant. All categories were similarly calculated, and none were significant at the 95% level.

POISSON DISTRIBUTION TEST—STEP 20A

PROJECT: *Roadside Obstacle*

PROJECT NO.: *CS-2*

EVALUATOR/DATE: *5/17/77 GHM*

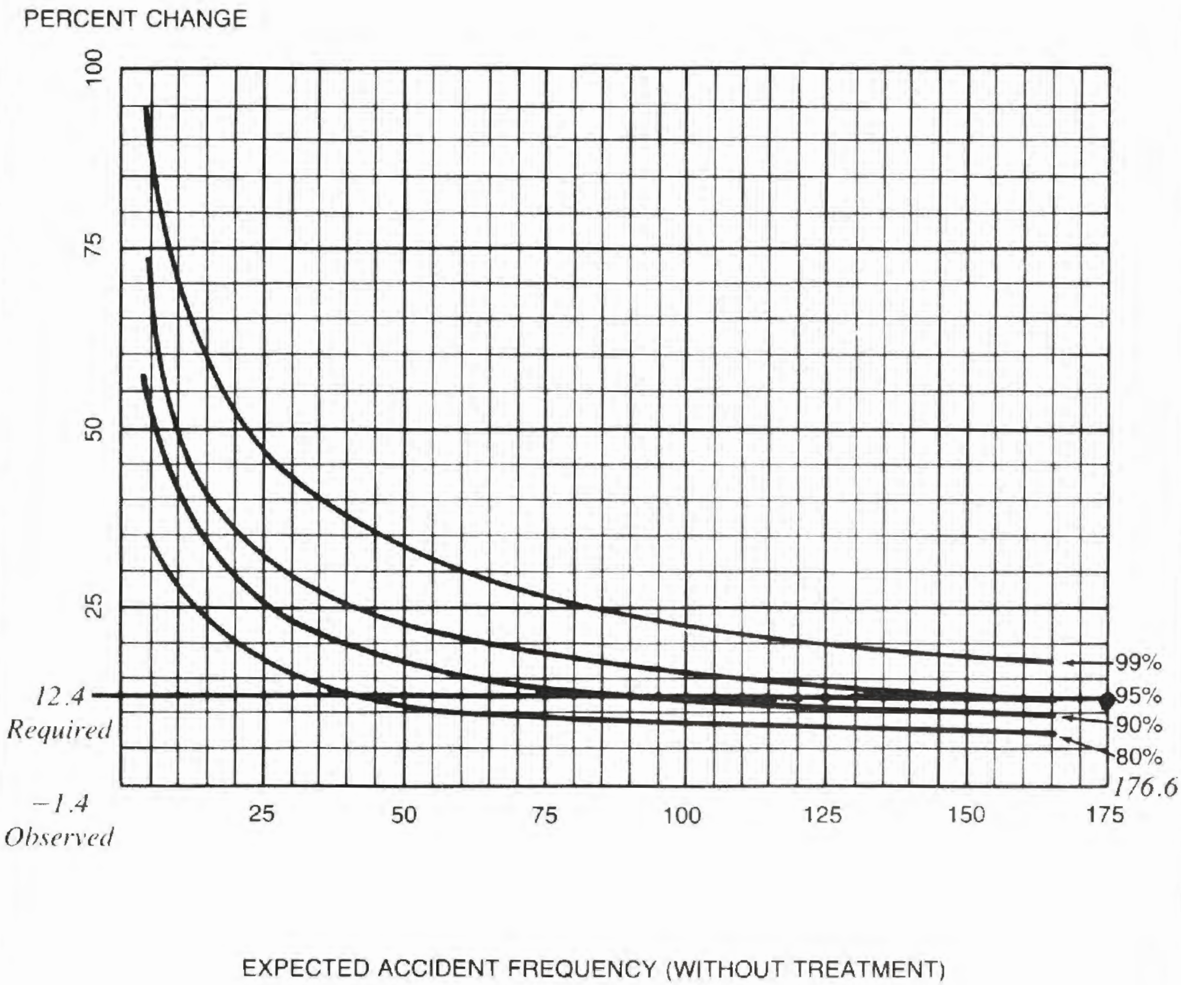
CONFIDENCE LEVEL: ☐ 80% ☐ 90% ☑ 95% ☐ Other ______

Accident Categories	"After" Frequency *3* Years		Percent Reduction		Significant For *3* yrs.
	Observed (A_{PF})	Expected (E_F)	Observed	Required	Yes or No*
From Step 1	From Step 19A		From Step 19A	From Curve At *95%* Confidence	
(Fundamental)					
Total Accidents	*179*	*176.6*	*−1.4*	*12*	*NO*
Fatal Accidents	*1*	*6.4*	*81.3*	*	*
Injury Accidents	*58*	*66.6*	*12.6*	*18*	*NO*
PDO Accidents	*120*	*104.0*	*−15.3*	*16*	*NO*
(Project Objectives)					
Fixed Object					
Total	*59*	*59.8*	*1.3*	*23*	*NO*
F&I	*24*	*29.9*	*20.0*	*29*	*NO*

*Too small to test

Exhibit 4-8.
Poisson Distribution Test Example

POISSON DISTRIBUTION



(Example for Total Accident Rate Change)

Exhibit 4-8.
(Continued)

STEP 20B. t Test

The recommended test to use to evaluate whether reductions in continuously distributed parametric MOE's, such as delay, lateral placement, passing distance, spot speeds and time through intersections, are significant is the t Test. The t Test is based on a sampling distribution referred to as the t distribution. The t distribution is the ratio of the mean (average) of a distribution of samples to the standard error of the sampling distribution. For example, if one were to take a number of samples of spot speed and find the average value of each sample, and then form a distribution of these average values, a distribution of samples of spot speeds would be derived. The average value of that distribution would be the mean of the distribution of samples of spot speeds. The standard error is a measure of the amount that the mean value of the statistic, in this case spot speed, may be expected to differ by chance from the true value of the statistic.

Since t is determined by the ratio of a statistic to its standard error, and the standard error is determined by the square root of the ratio of the standard deviation to the sample size minus one, t is sensitive to the number of observations in the sample. There are an infinite number of t distributions possible, from a sample size of 1 to infinity. However, when the number of observations exceeds 30, all t distributions are essentially the same.

When it can be shown that the probability of a difference between "Before" and "After" results occuring by chance is low (based on a preselected acceptance criterion), then it can be concluded that the improvement had a significant effect on the MOE. If, on the other hand, a difference can be shown to be attributed to chance, then it can be concluded that the improvement had no significant effect. In interpreting t, a quantity called degrees of freedom is used. "Degrees of freedom" is expressed as the number of observations minus the number of calculations used to estimate a particular statistic (sample size minus two).

Significance is determined by calculating the ratio of the difference in the MOE to the standard error and seeing whether this ratio falls above or below the t value needed for significance (for the appropriate degrees of freedom at the preselected confidence level). The t values are contained in a standard t table. If the ratio is equal to or greater than the value in the table, then it is concluded that the results are significant. If the ratio is less than the value in the table, the results are not significant.

A t Test example is presented in Exhibit 4-9. In this example, edge markings and speed advisories are applied to a rural two-lane road. The difference in "Before" and "After" mean speed is evaluated at an 80% confidence level (See Exhibit 3-8 for the "Before" data summary).

The mean "Before" speed (42.3 mph) and standard deviation (4.5 mph) and the mean "After" speed (39.6 mph) and standard deviation (4.5 mph) are inserted in the formula, yielding a t value of 5.8. "Degrees of freedom" are calculated as 375. This value is greater than 60 and is interpreted as "infinity." The critical t value (0.84) is found in the table on Sheet 2 at an 80% confidence level and infinite degrees of freedom. The calculated value (5.8) is compared to 0.84, and since it is larger, the results are significant.

t TEST—STEP 20B

PROJECT: *Spot Speed Study, Rural Two Lane Road*

PROJECT NO.: *EV-1*

EVALUATOR/DATE: *HL 8/17/80*

MOE: *Spot Speeds*

CONFIDENCE LEVEL: ☑ 80% ☐ 90% ☐ 95% ☐ ______ Other

"BEFORE" DATA—FROM STEP 13	"AFTER" DATA—FROM STEP 18
$\bar{X}_B$ = 42.3 MPH	$\bar{X}_A$ = 39.6 MPH
N_B = 186	N_A = 191
SD_B = 4.5 MPH	SD_A = 4.5 MPH

$$t = \frac{\bar{X}_B - \bar{X}_A}{\sqrt{\dfrac{N_B SD_B^2 + N_A SD_A^2}{N_B + N_A - 2}\left(\dfrac{N_B + N_A}{N_B N_A}\right)}} = \frac{42.3 - 39.6}{\sqrt{\dfrac{(186)(4.5)^2 + (191)(4.5)^2}{(186) + (191) - 2}\left(\dfrac{186 + 191}{186 \times 191}\right)}}$$

$$= \frac{2.7}{\sqrt{(20.238)(.0106)}} = \frac{2.7}{.463} = 5.83$$

DEGREES OF FREEDOM = $N_B + N_A - 2$ = 186 + 191 − 2 = 375

t VALUE (FROM TABLE) WITH ∞ DEGREES OF FREEDOM

AND 80 % CONFIDENCE LEVEL = 0.84.

COMPUTED t = 5.83

RESULTS: ☑ SIGNIFICANT ☐ NOT SIGNIFICANT

Exhibit 4-9.
t Test Example

t TEST

VALUES OF t

Degrees of Freedom	Levels of Confidence 80%	90%	95%	99%
1	1.38	3.08	6.31	31.82
2	1.06	1.89	2.92	7.00
3	0.98	1.64	2.35	4.54
4	0.94	1.53	2.13	3.75
5	0.92	1.48	2.02	3.37
6	0.91	1.44	1.94	3.14
7	0.90	1.42	1.90	3.00
8	0.89	1.40	1.86	2.90
9	0.88	1.40	1.83	2.82
10	0.88	1.37	1.81	2.76
11	0.88	1.36	1.80	2.72
12	0.87	1.36	1.78	2.68
13	0.87	1.35	1.77	2.65
14	0.87	1.35	1.76	2.62
15	0.87	1.34	1.75	2.60
16	0.87	1.34	1.75	2.58
17	0.87	1.33	1.74	2.58
18	0.86	1.33	1.73	2.55
19	0.86	1.33	1.73	2.54
20	0.86	1.33	1.73	2.53
21	0.86	1.32	1.72	2.52
22	0.86	1.32	1.72	2.51
23	0.86	1.32	1.71	2.50
24	0.86	1.32	1.71	2.49
25	0.86	1.32	1.71	2.49
26	0.86	1.32	1.71	2.48
27	0.86	1.31	1.70	2.47
28	0.86	1.31	1.70	2.47
29	0.85	1.31	1.70	2.46
30	0.85	1.31	1.70	2.46
40	0.85	1.30	1.68	2.42
50	0.85	1.30	1.67	2.39
60	0.85	1.29	1.66	2.36
∞	0.84	1.28	1.65	2.33

Exhibit 4-9. (Continued)

STEP 20C. F Test

The recommended test to use to determine whether a reduction in the variability (as measured by the standard deviation) of continuously distributed MOE's such as speed variance and lateral placement variance is the F Test. The F test is based on a sampling distribution referred to as the F distribution. F is the ratio of the two variances that correspond to the "Before" and "After" standard deviations. Variance is a measure of dispersion which indicates the extent to which each individual measure differs from any other measure in a given sample. It is obtained by squaring the standard deviation. The larger variance is compared to the smaller variance. If their ratio is 1.00, then the two variances are equal. The greater the difference in the variance, the larger the ratio. When this ratio exceeds a critical value for a given confidence level, then the difference in the variance is significant, indicating that the difference in the standard deviations is also significant.

To compute F, "Before" and "After" variances are each calculated. The larger variance, which could be either the "Before" or "After" variance, is then divided by the smaller variance and an F ratio obtained. F is interpreted in terms of degrees of freedom. For the F statistic, degrees of freedom equals the number of measures in the sample minus one. Degrees of freedom are calculated for the larger and smaller variances. The ratio is then looked up in an appropriate table for the preselected confidence level (F values do not exist for an 80% confidence level). The degrees of freedom of the larger variance (numerator or column) and the smaller variance (denominator or row) are used to find the critical F value for significance in the appropriate table. If the computed ratio is equal to or greater than the critical value in the table, the difference in variance is significant. If the value is less than the critical value, then the results are not significant.

An example of how F is calculated is shown in Exhibit 4-10. This example presents data for an urban signal timing project. Differences in speed variance are evaluated at the 90% confidence level (see Exhibit 3-9 for the "Before" data summary).

The "Before" standard deviation (4.18) is smaller than the "After" standard deviation, so it is squared and inserted in the denominator of the F formula. The "After" standard deviation squared becomes the numerator. The degrees of freedom for both variances is 13. However, given the table constraints, 12 degrees of freedom is used for the numerator. Using these values in the 90% Confidence Table, it is seen that an F of 2.10 is required for significance. The obtained value of 1.28 is, therefore, not significant.

F TEST—STEP 20C

PROJECT: *Caldon St.–Urban Arterial*

PROJECT NO.: *A3B*

EVALUATOR/DATE: *HL 4/25/80*

MOE: *Speed Variance*

CONFIDENCE LEVEL: ☑ 90% ☐ 95% ☐ ______ Other

"BEFORE" DATA-FROM STEP 13	"AFTER" DATA-FROM STEP 18
SD_B = *4.18*	SD_A = *4.70*
N_B = *14*	N_A = *14*

LARGER SD_L = *4.70* N_L = *14*

SMALLER SD_S = *4.18* N_S = *14*

LARGER VARIANCE = S_L^2 = (LARGER SD)^2 = (*4.70*)^2 = *22.17*

SMALLER VARIANCE = S_S^2 = (SMALLER SD)^2 = (*4.18*)^2 = *17.45*

$$F = \frac{S_L^2}{S_S^2} = \frac{22.17}{17.45} = 1.28$$

LARGER DEGREES OF FREEDOM = $N_L - 1$ = *13 (table uses 12)*

SMALLER DEGREES OF FREEDOM = $N_S - 1$ = *13*

F VALUE (FROM TABLE) WITH *12* DEGREES OF FREEDOM

FROM NUMERATOR (LARGER) AND *13* DEGREES OF FREEDOM

FROM DENOMINATOR (SMALLER) AND *90* % CONFIDENCE LEVEL = *2.10*.

COMPUTED F = *1.28*

RESULTS: ☐ SIGNIFICANT ☑ NOT SIGNIFICANT

Exhibit 4-10.
F Test Example

VALUES OF F (.10 LEVEL)

90% Confidence Level

f_1 = DEGREES OF FREEDOM FOR NUMERATOR

f_2 = DEGREES OF FREEDOM FOR DENOMINATOR

f_2	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.20	60.70	61.22	61.74	62.00	62.26	62.53	62.79	63.06	63.33
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48	9.49
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.20	5.18	5.18	5.17	5.16	5.15	5.14	5.13
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.79	3.78	3.76
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14	3.12	3.10
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.90	2.87	2.84	2.82	2.80	2.78	2.76	2.74	2.72
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.67	2.63	2.59	2.58	2.56	2.54	2.51	2.40	2.47
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.34	2.32	2.29
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.38	2.34	2.30	2.28	2.25	2.23	2.21	2.18	2.16
10	3.28	3.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.28	2.24	2.20	2.18	2.16	2.13	2.11	2.08	2.06
11	3.23	2.86	2.68	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.03	2.00	1.97
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.15	2.10	2.06	2.04	2.01	1.99	1.96	1.93	1.90
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.90	1.88	1.85
14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10	2.05	2.01	1.96	1.94	1.91	1.89	1.86	1.83	1.80
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	2.02	1.97	1.92	1.90	1.87	1.85	1.82	1.79	1.76
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03	1.99	1.94	1.89	1.87	1.84	1.81	1.78	1.75	1.72
17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00	1.96	1.91	1.86	1.84	1.81	1.78	1.75	1.72	1.69
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98	1.93	1.89	1.84	1.81	1.78	1.75	1.72	1.69	1.66
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.9[illegible]	1.96	1.91	1.86	1.81	1.79	1.76	1.73	1.70	1.67	1.63
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.89	1.84	1.79	1.77	1.74	1.71	1.68	1.64	1.61
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92	1.88	1.83	1.78	1.75	1.72	1.69	1.66	1.62	1.59
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.64	1.60	1.57
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89	1.84	1.80	1.74	1.72	1.69	1.66	1.62	1.59	1.55
24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.61	1.57	1.53
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87	1.82	1.77	1.72	1.69	1.66	1.63	1.59	1.56	1.52
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.81	1.76	1.71	1.68	1.65	1.61	1.58	1.54	1.50
27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87	1.85	1.80	1.75	1.70	1.67	1.64	1.60	1.57	1.53	1.49
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.79	1.74	1.69	1.66	1.63	1.59	1.56	1.52	1.48
29	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83	1.78	1.73	1.68	1.65	1.62	1.58	1.55	1.51	1.47
30	2.88	2.49	2.26	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.77	1.72	1.67	1.64	1.61	1.57	1.54	1.50	1.46
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.47	1.42	1.38
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.66	1.60	1.54	1.51	1.48	1.44	1.40	1.35	1.29
120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.60	1.54	1.48	1.45	1.41	1.37	1.32	1.26	1.19
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.55	1.49	1.42	1.38	1.34	1.30	1.24	1.17	1.00

VALUES OF F (.05 LEVEL)

95% Confidence Level

f_1 = DEGREES OF FREEDOM FOR NUMERATOR

f_2 = DEGREES OF FREEDOM FOR DENOMINATOR

f_2	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161.45	199.50	215.71	224.56	230.16	233.99	236.77	238.88	240.54	241.88	243.91	245.95	248.01	249.05	250.09	251.14	252.20	253.25	254.32
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.00	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.48	4.43	4.40	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.84	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.00	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.74	1.79	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.48	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.36	1.23
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.62	1.46	1.39	1.32	1.22	1.00

Exhibit 4-10.
(Continued)

STEP 20D. Z Test

The recommended test to use to determine whether reductions in traffic conflicts and erratic maneuvers (where the number of observations equals or is greater than 30) is significant is the Z Test. The Z Test is a test which is based on the assumption that the distribution of the MOE approximates a normal, bell-shaped one. If the data is converted to standard scores (a standard score is one using as its units the standard deviation of the MOE) called Z scores, then the critical value for significance can be determined directly from the normal distribution for any appropriate confidence level.

Since the number of events must equal or exceed 30 to use the Z Test (Chi Square, described in Step 20E, is used for less than 30 events), it is not necessary to compute degrees of freedom. The critical value is obtained directly from the Z table for any confidence level of from 80% to 99%. Once Z is computed, it is evaluated in the same manner that t and F are interpreted. That is, if the computed Z value is equal to or greater than the critical table value, then the results are significant; if not, then the results are not significant.

A Z Test example is shown in Exhibit 4-11. This example is for a diagrammatic guide sign improvement at a left hand exit. Differences in erratic maneuvers are evaluated at the 80% level.

The "Before" erratic maneuvers (71/1,000) and "After" erratic maneuvers (30/1,000) are used in the formula to compute a Z of 3.2. This value, at an 80% confidence level, exceeds the 0.842 required for significance. The improvement is, therefore, significant.

Z TEST—STEP 20D

PROJECT: *Diagrammatic Guide Signs –I111 & US333*

PROJECT NO.: *B12*

EVALUATOR/DATE: *HL 9/4/79*

MOE: *Erratic Maneuvers*

CONFIDENCE LEVEL: ☑ 80% ☐ 90% ☐ 95% ☐ ________ Other

n_1 = "Before" erratic maneuvers or traffic conflicts = *71/1000*

n_2 = "After" erratic maneuvers or traffic conflicts = *39/1000*

N_1 = Traffic Volume – "Before" = *1000*

N_2 = Traffic Volume – "After" = *1000*

$n = n_1 + n_2$ = *71 + 39 = 110*

$N = N_1 + N_2$ = *2000*

$$Z = \frac{N_2 n_1 - N_1 n_2}{\sqrt{\dfrac{N_1 N_2 n (N - n)}{N}}} \qquad \frac{71 \times 10^3 - 39 \times 10^3}{\sqrt{\dfrac{10^3 \times 10^3 \times (1.1 \times 10^2)(2 \times 10^3 - 110)}{2 \times 10^3}}}$$

$$= \frac{3.2 \times 10^4}{\sqrt{1.04 \times 10^8}} = 3.2$$

VALUES OF Z	CONFIDENCE LEVEL			
	80% = 0.842	90% = 1.262	95% = 1.645	99% = 1.960

Z VALUE (FROM TABLE) WITH *80* %

CONFIDENCE LEVEL = *0.842*

COMPUTED Z = *3.2*

RESULTS: ☑ SIGNIFICANT ☐ NOT SIGNIFICANT

Exhibit 4-11.
Z Test Example

STEP 20E. <u>Chi Square Test</u>

Chi Square is the recommended test to use to determine whether reductions are significant for: Brake applications, conflicts with small sample size (less than 30); delay (when a t Test cannot be used); encroachments; erratic maneuvers with small sample size (less than 30); lane changes; lateral placement (when scored categorically); merges; passing frequency or type; points of entry; speed profiles; and any other MOE which is scored by category or when sample sizes are small or when parametric assumptions cannot be met. As such, Chi Square is a general purpose test with wide application. In fact, Chi Square could be used to test for significance for any MOE.

Chi Square provides a means of estimating whether the obtained result differs from what would be expected if no improvement were made to such a degree that nonchance factors, i.e., the improvement, can be judged to be operative. Thus, if it is possible to get a large enough deviation based on a predetermined confidence level, the results are deemed significant. This is determined by calculating the Chi Square value and comparing it to the critical value in a Chi Square table. Chi Square is evaluated at a given confidence level for a calculated number of degrees of freedom. Significance is achieved when the calculated Chi Square value exceeds the critical table value.

To use Chi Square, the data must be tabulated. In general, the data will be tabulated in a contingency table. A contingency table is a two-way table showing the frequency of occurrence of the events or categories as indicated by the horizontal rows and the "Before" and "After" results as indicated by the vertical columns. The contingency table also serves to compute the degrees of freedom. When Chi Square is computed from a one-dimensional set of n categories, the degrees of freedom equals the number of categories minus one. When Chi Square is computed from a cross-tabulation of r rows and c columns, the degrees of freedom equals rows minus one times columns minus one. The latter case is one that generally will not be used, since the basic Chi Square calculations will be for "Before" and "After" data in a contingency table. Note that the degrees of freedom are calculated on the basis of the number of categories rather than data points.

In Exhibit 4-12, a Chi Square example is presented for an urban intersection project. The project involved adding a right turn arrow and pedestrian signals. The MOE was traffic conflicts, which were tested at the 80% confidence level.

The "Before" (fo) and "After" (fh) values are inserted in the table and differences are determined and squared. The squared values are divided by the "After" values and this column is summed to yield a Chi Square value of 17.30. Note that volume corrections must be used to insure that the total of fo and fh are zero. It is also important to note that fo-fh must equal zero. The degrees of freedom is determined to be 7. Referring to the table on page 2, with 7 degrees of freedom and an 80% confidence level, a Chi Square of 7.80 or greater is required. Since the computed value of Chi Square is 17.30, the results are significant.

CHI SQUARE—STEP 20E

PROJECT: *Urban Interchange, 6th & D*

PROJECT NO.: *AYZ*

EVALUATOR/DATE: *HL 4/5/80*

MOE: *Traffic Conflicts*

Event	"Before" fo	"After" fh	fo−fh	$(fo-fh)^2$	$\frac{(fo-fh)^2}{fh}$
Left turn	*20*	*19*	*+ 1*	*1*	*1/19 = .05*
Weave	*20*	*18*	*+ 2*	*4*	*4/18 = .22*
Rear-end–amber	*24*	*20*	*+ 4*	*16*	*16/20 = 0.80*
– Thru lane	*16*	*15*	*+ 1*	*1*	*1/15 = 0.07*
– Left turn	*16*	*16*	*0*	*0*	*0/16 = 0*
– Right turn	*14*	*6*	*+ 8*	*64*	*64/6 = 10.69*
– Pedestrian	*10*	*5*	*+ 5*	*25*	*25/5 = 5.00*
Normal maneuvers	*880*	*901*	*−21*	*441*	*441/901 = 0.49*
TOTAL (Σ)	*1000*	*1000*	*0*	*—*	*17.30*

DEGREES OF FREEDOM = NUMBER OF EVENTS −1 = *8 − 1 = 7*

CONFIDENCE LEVEL: ☑ 80% (.20) ☐ 90% (.10) ☐ 95% (.05) ☐ ______ Other

CHI SQUARE VALUE (FROM TABLE) WITH *7* DEGREES OF FREEDOM AND *80* % CONFIDENCE LEVEL = *9.80*.

COMPUTED CHI SQUARE = *17.30*

RESULTS: ☑ SIGNIFICANT ☐ NOT SIGNIFICANT

Exhibit 4-12.
CHI Square Example

CHI SQUARE

CHI SQUARE

df	Levels of Confidence 80%	90%	95%	99%
1	1.64	2.71	3.84	6.64
2	3.22	4.61	5.99	9.21
3	4.64	6.25	7.82	11.34
4	5.99	7.78	9.49	13.28
5	7.29	9.24	11.07	15.09
6	8.56	10.65	12.59	16.81
7	9.80	12.02	14.07	18.48
8	11.03	13.36	15.51	20.09
9	12.24	14.68	16.92	21.67
10	13.44	15.99	18.31	23.21
11	14.63	17.26	19.68	24.72
12	15.81	18.55	21.03	26.22
13	16.99	19.81	22.36	27.69
14	18.15	21.06	23.68	29.14
15	19.31	22.31	25.00	30.58
16	20.47	23.54	26.30	32.00
17	21.62	24.77	27.59	33.41
18	22.76	25.99	28.87	34.80
19	23.90	27.20	30.14	36.19
20	25.04	28.41	31.41	37.57
21	26.17	29.62	32.67	38.93
22	27.30	30.81	33.92	40.29
23	28.43	32.01	35.17	41.64
24	29.55	33.20	36.42	42.98
25	30.68	34.38	37.65	44.31
26	31.80	35.56	38.88	45.64
27	32.91	36.74	40.11	46.96
28	34.03	37.92	41.34	48.28
29	35.14	39.09	42.56	49.59
30	36.26	40.26	43.77	50.89

Exhibit 4-12.
(Continued)

STEP 21. Determine Practical Significance

During this step, all aspects of the evaluation are assessed in order to determine whether the improvement has achieved the objectives, or in cases where the objectives have not been achieved, why they were not met. In addition, evaluation results are further interpreted to obtain an indication of the degree to which objectives have been achieved. Finally, consideration is given as to what course of action to follow on the basis of the project results.

The outcome of the statistical tests yields the primary indication of the achievement of objectives. Obviously, results that are statistically significant are most likely to have achieved the project's objectives. Once it has been determined that the results are significant at a preselected level, it is acceptable to test the data for significance at a higher level. Since the higher the confidence level tested to, the more confidence that the results are not due to chance, it may be useful to test at a higher level during this step. Although achievement of significance does not automatically assure that the results are due to the improvement, close adherence to the steps in this procedure as well as statistical significance greatly enchances the conclusion that the improvement has achieved its objectives.

Further evidence of the degree to which the project objectives have been achieved can be obtained by examining the size and assesing the meaningfulness of the differences. This is accomplished by comparing the magnitude of the differences between "Before" and "After" MOE's to determine the extent to which the problem was reduced, and by assessing whether this reduction has any "real-world" significance. While any significant reduction in accidents is meaningful, there are a number of traffic and system performance MOE's where this determination is not as clear cut. Often small differences with little practical importance may achieve statistical significance, particularly if the sample is large. For example, a 1/2-inch shift in lateral position or a 1 mph reduction in spot speed may be statistically significant, while not being practically significant.

Practical significance is, thus, an important consideration in the overall project assessment. Results must be meaningful and useful. To determine practical significance often requires engineering judgment in assessing the size of the change vis a vis a number of factors such as the cost of the project and the ideal speed and path required. Usually, close adherence to a systematic procedure of problem analysis and solution development helps assure practical significance.

Another assessment that can be made, data permitting, is an economic assessment of the cost-benefit, cost-effectiveness, or energy savings of the improvement. This should only be considered if the results have achieved statistical significance, and when a monetary value can be placed on the improvement. This often limits an economic analysis to accident MOE's and time-related MOE's such as delay. The reader is referred to Datta et al. (1978) for suitable procedures, or to a standard highway engineering economy text, such as Winfrey (1969) or AASHTO (1977). Energy efficiency can be assessed using the methods contained in Dale (1980).

When the results are found to be both statistically and practically significant, there are a number of possible courses of action to follow. The most obvious is to leave the improvement in place. In addition, if there is a State or local file of successful improvements, the improvements should be incorporated into the file for future use. If there are similar problem locations, the improvement will be a candidate for immediate implementation. Finally, a report should be prepared outlining the results of the evaluation.

In cases where the results were not statistically and/or practically significant, project personnel should try to assess all aspects of the evaluation to see why the project objectives were not achieved. There is always the possibility that the improvement was not properly designed or applied, or that the evaluation was not properaly conducted. There may have been an intervening change that invalidated the results or some other factor that is not readily apparent. If negative results were attributable to a flaw in the evaluation, consideration may be given to rerunning it. It may only be necessary to make minor changes and partially reevaluate or, in extreme cases, the project may have to start from scratch.

Figure 4-4 shows the Step 21 flow.

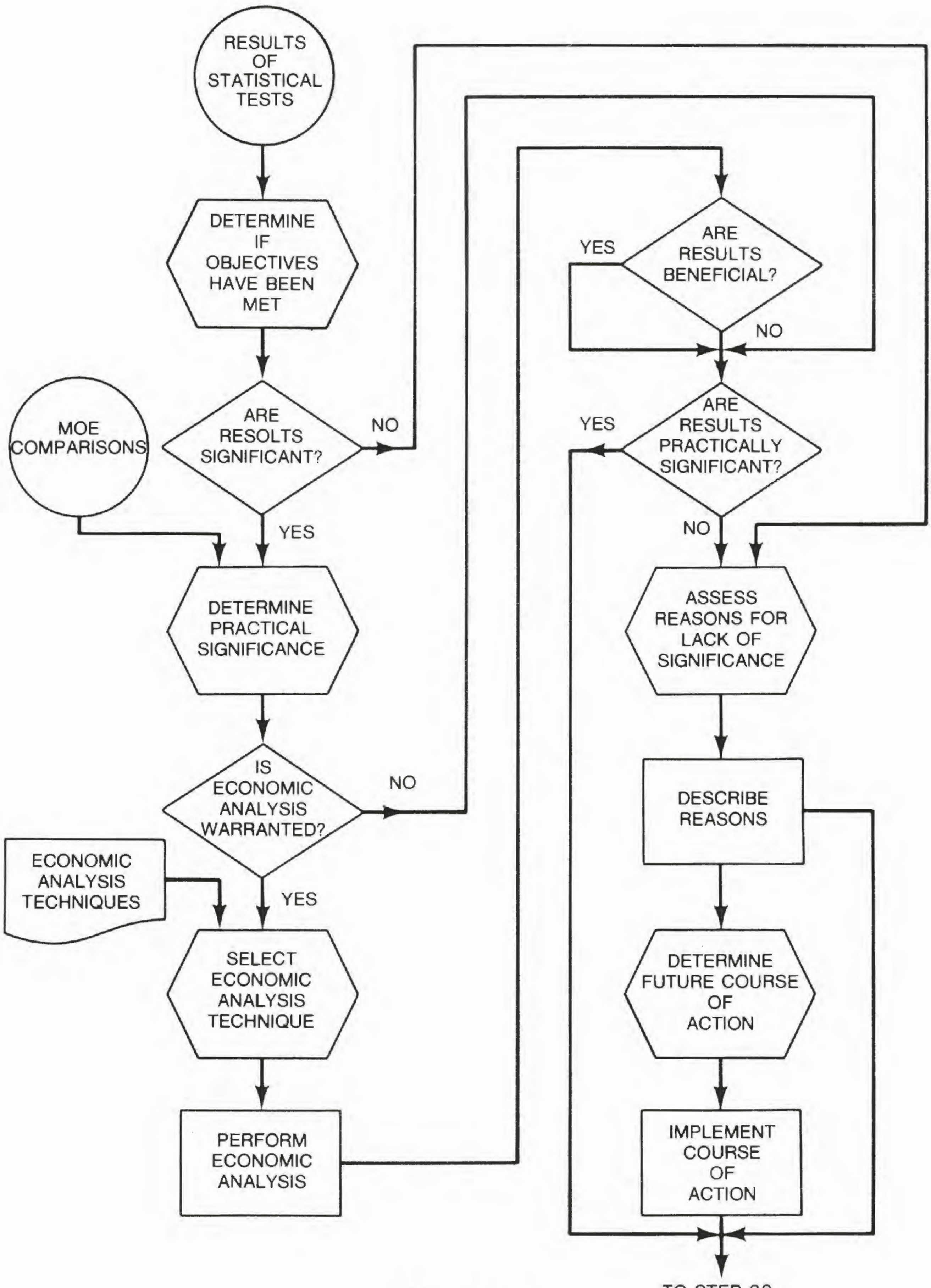


Figure 4-4.
Flow, Step 21 - Determine Practical Significance

STEP 22. Report Findings and Recommendations

The final step in the evaluation is to gather information from all steps and prepare a final report. The report provides information about the effectiveness of various traffic operations, safety, and Positive Guidance improvements. It also serves as valuable documentation for future activities.

Since evaluation reporting is used as feedback to improve the decisionmaking process, the report should document all relevant aspects of the improvement development, evaluation activities, and recommendations and conclusions regarding present implementation and future applications. It should contain sufficient details, including site and engineering diagrams, to enable others to assess the applicability of the improvement to their particular problem location. All reports should provide graphic presentations of "Before" and "After" information system treatments, statistical facts and figures, and critical results.

All formal reports should follow a standardized format. An example of a suitable format is shown in Exhibit 4-13. The extent to which the report is published and distributed is often a function of the State or local jurisdiction's procedures as well as the overall importance of the project.

The Step 22 flow is shown in Figure 4-5.

REPORT FORMAT

1. Executive Summary:

   Summary of project performance;
   Summary of successes, failures, unexpected impacts, and probable causes;
   Recommendations for project improvement;
   Quantifiable support for conclusions.

2. Introduction:

   Project name;
   Project overview and improvement type(s);
   Funding level and period;
   Key project personnel.

3. Identification and Problem Discussion:

   Site diagrams;
   Problem identification;
   Problem discussion;
   Project appropriateness discussion;
   Opinions.

4. Administrative Evaluation of the Project:

   Initiated and completed activities;
   Personnel involved;
   Time schedule;
   Costs.

5. Project Effectiveness Evaluation:

   Methods used, including
   Evaluation Design (i.e., purposes, objectives, MOE, Evaluation Plan, etc.);
   Data collected;
   Data collection and reduction procedures used;
   Detailed project results relative to achievement of objectives;
   Detailed project impact statement;
   Problems encountered in the overall evaluation.

6. Conclusions and Recommendations.

Exhibit 4-13. Evaluation Report Format

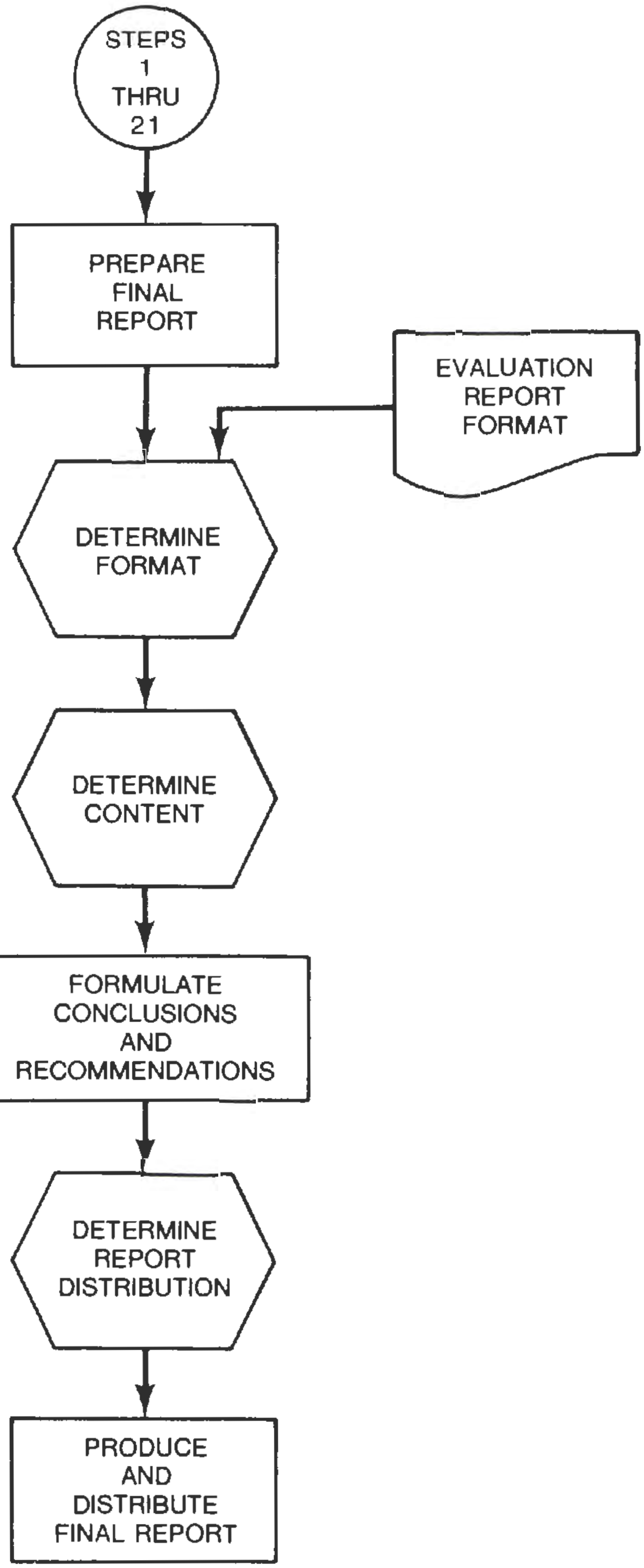


Figure 4-5.
Flow, Step 22 - Report Findings and Recommendations

GLOSSARY OF TERMS

Acclimation - The period of time needed by repeat drivers to become accustomed to the improvement. The time period between the application of an improvement and the collection of "After" data.

Actuarial - A data collection method based on historical records that states relationships in terms of frequency of occurrence.

"After" Data - MOE data collected following application of an improvement and suitable implementation and acclimation period. "After" data are used in comparison with MOE data collected prior to an improvement to test for significance.

Ameliorate - To improve or make better. While problems often are not "solved," improvements can lessen their adverse safety or operational effects.

Appropriateness - The suitability of a project's objectives and the measures used to evaluate an improvement.

Arithmetic Summaries - A way to summarize data in terms of several quantitative measures, generally relating to central tendency and dispersion.

Assumptions (Test) - The logic underlying tests of significance. Parametric tests assume that the data are distributed in accordance with the normal distribution. Nonparametric tests are distribution free.

"Before" Data - MOE data collected prior to the application of an improvement. "Before" data are compared with "After" data to test for significance.

Binomial Distribution - An algebraic expression taking the form (p + q) or (p - q). A binomial distribution is the binomial raised to a power $(p + q)^n$. The Poisson distribution is a special case of the binomial distribution.

Categorical - Data that is assigned to classifications based on qualitative rather than quantitative differences.

Central Tendency - A value calculated from a set of randomly drawn measures which represents the typical value of the data set.

Chi Square - A general purpose nonparametric test of significance for use with categorical data. Chi Square estimates whether an obtained distribution is sufficiently different from an expected distribution to indicate that nonchance factors are operative.

Comparability - Insuring that all "Before" and "After" data collection conditions, equipment, personnel and procedures are essentially the same so that the only aspect of the site to change is the application of an improvement.

Concomitant - A benefit that accompanies another benefit, with or without a causal relationship existing.

Conflict - An evasive action (as evidenced by a brake light application or lane change) or traffic violation (as evidenced by an infraction of an existing traffic regulation).

Contingency Plans - Alternative plans made in case such unexpected things as traffic accidents, adverse climatological conditions, equipment failure etc., occur in the course of data collection.

Contingency Table - A two-way table showing the frequency of occurrence of classes or categories (horizontal rows) by "Before" and "After" results (vertical columns).

Continuous Variable - A quantitative attribute, measured by a numerical scale with a zero point, and capable of being infinitely subdivided.

Control Site - A site that exhibits all the attributes of the site being evaluated. The control site receives no improvements, thereby enabling a determination to be made whether there are changes over time occurring when "Before" and "After" data collected at the control site are compared.

Correlation - The relationship between two variables such that changes in one variable is accompanied by a corresponding change (positive or negative) in the other.

Cost-Benefit Analysis - An economic evaluation where inputs are measured in dollar costs, and outputs are measured in terms of economic benefits of an improvement compared to its costs.

Cost-Effectiveness Analysis - An economic evaluation where inputs are measured in terms of project effectiveness and outputs are measured in terms of the cost of achieving one unit of the measure of effectiveness.

Critical Ratio - The ratio of a statistic to its standard error. When the results of a statistical test exceed this value (determined by the level of significance and the degrees of freedom), the results are significant.

Data Set - All the data pertaining to the project, or collected during a single data collection period.

Degrees of Freedom - The number of observations minus the number of calculations used to estimate a particular statistic.

Dependant Variable - The measure of effectiveness. A variable whose changes are tested as being due to changes in one or more other variables (called the independent variable(s)).

Descriptive Statistics - Statistics used to organize and describe the sample from which they were derived in a meaningful manner using tables, graphs, and/or arithmetic summaries.

Diagnostic Measures - Those measures of effectiveness which also serve to diagnose the causes of the problem.

Distribution - A systematic way to group data in a meaningful way in terms of frequency of occurrence.

Dispersion - The extent to which measures differ from the central value.

Empirical - Based on observation and investigation.

Erratic Maneuvers - Deviations from an idealized trace through an interchange or intersection, given a particular destination.

Evaluation Plan - A comprehensive framework outlining all aspects of the evaluation.

Experimental Design - The analytical framework for determining the effect of the improvement on the measures of effectiveness.

Exposure - The quantity of vehicles, vehicle-miles of travel, or other volume and/or time related factors which measure the degree of vehicular exposure to a particular situation.

F Test - A parametric test used to determine whether reductions in variability, as measured by variance, are significant.

Feasibility - The ability of an evaluator to perform a particular operation or collect a certain measure of effectiveness. Feasibility rests in the availability of resources and the characteristics of the site.

Frequency - The number of occurrences of a particular measure of effectiveness over a given time period (for example, accidents per year).

Frequency Distribution - The number of times a given value of a variable occur. A frequency distribution can be graphic or in tabular form.

Generalizability - The ability to reach a conclusion about a whole class on the basis of limited experience with a limited sample of that class.

Grouping - The process of combining scores or measures into categories or ranks.

Hazard Potential - Conditions or situations which are conducive to future accidents.

Hazardous Maneuvers - Conflicts or erratic maneuvers with a high risk of accident involvement.

Independent Variable - Characteristics of the environment or drivers. Changes are not dependent on changes in the measures of effectiveness.

Inferential Statistics - Statistical procedures which enable evaluators to go beyond the sample and estimate how close the sample is to the population. Inferential statistics allow for forecasting and decisionmaking.

Interval Scale - A measurement scale where the distance between each successive point is equal.

Large Sample - A sample with 30 or more observations.

Level of Confidence - The degree of confidence, in percent, that the difference between "Before" and "After" data was not due to chance. The term is analogous to level of significance.

Level of Significance - The probability, in decimals, that the difference between "Before" and "After" data was due to chance. The term is analogous to level of confidence.

Mean - A measure of central tendency. The mean is the average value and is obtained by summing all measures and dividing by the number of observations.

Measures of Effectiveness (MOE) - Those accident, system performance or traffic performance measures used to evaluate the effectiveness of an improvement.

Median - The score in a distribution which has half the scores below and half the scores above it.

Mode - The most common score, where a frequency distribution curve peaks (there can be more than one mode if a curve has several peaks).

Nonparametric - A test which makes no assumptions about the underlying population distribution.

Normal Distribution - A bell shaped frequency distribution which describes the expected or usual values of a variable when observations are infinite and the variation is subject to the laws of chance. The mean, median, and mode all coincide in a normal distribution.

Novelty Affect - Transitory changes in behavior of repeat drivers caused by the novelty of an improvement.

Null Hypothesis - That which is tested in a test of significance. The usual null hypothesis is that the results are due to chance. If this can be disproved, then the null hypothesis is rejected and the results are judged statistically significant.

Obtrusiveness - A situation where the equipment and/or personnel used to collect data are visible and potentially change driver behavior.

Operational Definition - A definition of a measure of effectiveness in precise, behavioral terms. The operations or procedures used to distinguish the MOE from other measures forms the operational definition.

Parameter - An attribute of the overall population.

Parametric - A test where the underlying assumption is that the population from which the sample was drawn is normally distributed, and that the measure is continuous and interval in nature.

Percentile - A score which divides a ranked distribution into groups or parts. There are 99 possible percentiles, ranging from 1st to 99th. The score represents a percentage of the total number of cases, such as 85th percentile speed.

Pilot Testing - A brief and simplified test to try out equipment, methods, procedures, forms, etc.

Poisson Distribution Test - A test for the significance of accident and other rare event data based on a special distribution called the Poisson distribution. The test determines if the "Before" and "After" accidents come from the same distribution.

Population - The entire amount of measures that define the universe of interest, i.e., all the speeds, every driver, etc.

Practical Significance - The extent to which obtained results are meaningful and useful.

Precision of Measurement - The amount of sampling error permitted. The more precise the measure, the larger the sample size.

Qualitative - A measurement which assigns a position of greater than or less than or the existence of a quantity (e.g., erratic maneuver) without any additional quantification.

Quantitative - A measure of a quantity that assigns a numerical magnitude (e.g., miles per hour) to the quantity.

Random - By chance. Selecting a sample in such a way that every vehicle or driver in the population has an equal and independent chance of being included.

Ranking - Arranging measures in an order according to a specific criteria so that each is greater (or smaller) than any other.

Range - A measure of dispersion encompassing the lowest and highest scores.

Rare Events - An occurrence such as an accident with a very low probability.

Rate - The number of occurrences of a particular measure of effectiveness during a given time period divided by the degree of vehicular exposure over the same time period (for example, erratic maneuvers per 1000 vehicles).

Regression, Linear - Computing the most likely value of one variable from the known value of another by obtaining the line which best fits the means of the columns and rows in a correlation chart.

Regression Toward the Mean - A situation where, when two measures are not highly associated, then unusually high scores measured one time will tend to be associated with average (mean) scores the next time.

Reliability - The ability of the measuring instrument to obtain similar results each time the measurement is made.

Repeatability - The ability of a measurement to be taken in the same or similar way at successive times.

Replicate - To repeat an evaluation with no changes in any essential aspect such as conditions, measures, equipment, and improvements.

Representativeness - The ability of the conditions under which the data are collected to replicate the problem conditions.

Sample - A portion of the population which is representative of the population.

Sampling - The process of drawing a representative sample.

Sampling Distribution - A frequency distribution of samples.

Selection Bias - Choosing vehicles or drivers in such a manner as to favor a particular group or to lead to a nonrepresentative sample.

Significance Testing - The process whereby differences "Before" and "After" an improvement are compared to determine whether the result could have occurred by chance.

Small Sample - A sample with less than 30 observations.

Spot Improvement - An improvement applied to a location such as an interchange or intersection.

Spurious - Not a real result, but one which is similar to what might be expected and which can be attributed to identifiable events or causes.

Stability - The ability of a measurement to remain the same whenever conditions are repeated.

Standard Deviation - A measure of variability (dispersion) indicating how spread out the data area. The standard deviation is the square root of the sum of the squared differences between a score and the mean, divided by the number of scores in the sample.

Standard Error - An estimate, calculated from the standard deviation, of the sampling error of the obtained statistic (such as the mean) from the true population value.

Standard Score - A score using as its units the standard deviation of the MOE.

Statistical Significance - A judgment established on the basis of the probability (level of significance) of results of an evaluation not being due to chance.

Surrogate - A measure such as erratic maneuvers and conflicts used as a substitute for accidents.

System Performance - Project objectives designed to improve operations of the system (e.g., traffic flow, delay). Also characterizes the measure of effectiveness used to measure the improvement.

t Test - A test of significance for continuously distributed interval scale parametric MOE's. t is based on the ratio of the mean of a distribution of samples to the standard error of the sampling distribution.

Traffic Performance - Project objectives designed to improve speed and path of vehicles and traffic (e.g., speed, lateral placement, erratic maneuver, conflicts, etc.). Also characterizes the measure of effectiveness used to measure the improvement.

Type I Error - Rejecting the null hypothesis when it is true. That is, saying the results are statistically significant when they are due to chance. The level of significance is the probability of committing a Type I error.

Type II Error - Accepting the null hypothesis when it is false. That is, saying the results are due to chance when they are actually statistically significant.

Validity - Proper reasoning used to arrive at a result. Also, the correctness of a measure in measuring the attribute it is supposed to measure.

Variable - The attribute used to define an aspect of the population of interest that distinguishes individuals or vehicles within the population.

Variance - A measure of dispersion which indicates the extent to which each individual measure differs from another measure in a sample. Variance is obtained by squaring the standard deviation.

Z Test - A test of significance based on large sample parametric data, where the assumption is that the MOE is normally distributed. Data are converted to standard scores and a critical ratio is calculated.

REFERENCES AND BIBLIOGRAPHY

AASHTO. A Manual on User Benefit Analysis of Highway and Bus-Transit Improvements. American Association of State Highway and Transportation Officials, Washington, D.C. 1977.

Bauerwald, J. E. (ED), Transportation and Traffic Engineering Handbook. Institute of Transportation Engineers, Arlington, Virginia, 1976.

Box, P. C. and Oppenlander, J. C. Manual of Traffic Engineering Studies. Fourth Edition. Institute of Transportation Engineers, Arlington, Virginia, 1976.

Dale, C. W. Procedure for Estimating Highway User Costs, Fuel Consumption and Air Pollution. FHWA, Washington, D.C. 1980.

Datta, T. K., Perkins, D. D. and Taylor, W. C. Evaluation of Highway Safety Projects. FHWA, Washington, D.C. 1978.

Dixon, W. J. and Massey, F. J. Introduction to Statistical Analysis. Third Edition. McGraw-Hill, New York, 1969.

Everall, P. F. Urban Freeway Surveillance and Control. FHWA, Washington, D.C. November 1972.

Gerlough, D. L. and Barnes, F. C. Poisson and Other Distributions in Traffic. The Eno Foundation, Connecticut, 1971.

Greenshields, B. D. and Weida, F. M. Statistics with Applications to Highway Traffic Analysis. The Eno Foundation, Connecticut, 1952.

Griffin, L. I., Powers, B. and Mullen, C. Impediments to the Evaluation of Highway Safety Programs. University of North Carolina, Chapel Hill, North Carolina, June 1975.

Guilford, J. P. Fundamental Statistics in Education and Psychology. McGraw-Hill, New York, 1956.

Hanscom, F. R. and Berger, W. G. Motorist Response to Highway Guide Signs. BioTechnology, Inc., Falls Church, Virginia. January 31, 1976.

Highway Research Board. Special Report 87. Highway Capacity Manual. Highway Research Board, Washington, D.C., 1965.

Hostetter, R. S. Collection of Data. FHWA Report FHWA-TO-80-2. IN PRESS.

Laughland, J. C., Haefner, L. E., Hall, J. W. and Clough, D. R. NCHRP Report 162. Methods for Evaluating Highway Safety Improvements. Transportation Research Board, Washington, D.C., 1975.

Mast, T. M. and Kolsrud, G. S. Diagrammatic Guide Signs for Use on Controlled Access Highways. Report No. FHWA-RD-73-21. December 1972, FHWA, Washington, D.C.

Midwest Research Institute. Application of Traffic Conflicts Analysis at Intersections. NCHRP Project 17-3. Transportation Research Board, Washington, D.C. IN PRESS.

Reilly, W. R., Gardner, C. C. and Kell, J. H. A Technique for Measurement of Delay at Intersections. Report No. FHWA-RD-76-135. FHWA, Washington, D.C., September 1976.

Siegal, S. Nonparametric Statistics for the Behavioral Sciences. McGraw-Hill, New York, 1956.

Taylor, J. I., McGee, H. W., Seguin, E. L. and Hostetter, R. S. NCHRP Report 130. Roadway Delineation Systems. Highway Research Board, Washington, D.C. 1972.

Votaw, D. F. and Levinson, H. S. Elementary Sampling for Traffic Engineers. The Eno Foundation, Connecticut, 1962.

Winfrey, R. Economic Analysis for Highways. International Textbook Company, Scranton, Pennsylvania, 1969.

APPENDIX A - STATISTICAL FACTORS

This appendix is designed to briefly touch on relevant statistical factors. Those users interested in more details should consult standard statistical texts (e.g., Dixon and Massey, 1969; Guilford, 1956; Siegal, 1956).

Descriptive Statistics - Important descriptive statistical factors include: Data types; distributions; and arithmetic summations.

Data Type -The attribute used to define an aspect of the population of interest that distinguishes individuals or vehicles within the population is called a variable. Variables are either quantitative if numbers can be assigned or qualitative when no numerical values can be assigned. Quantitative variables are usually measured using numerical, equal interval, continuous scales with a zero point. For this reason such data are classified as interval and continuous. Examples of variables that yield continuous data include speed, volume and traveltime. Another important quantifiable variable is accidents. Unlike variables measured using interval scales, accidents are random (occurring by chance) in nature and are often referred to as yielding rare event data. Qualitative variables are measured by assigning individuals into mutually exclusive categories. Data generated by these variables are referred to as categorical or count data. Examples include counts of vehicle types, erratic maneuvers, and brake light applications.

Distribution - Raw data conveys little useful information until it is tabulated and displayed in a meaningful way. Categorical or count data are generally displayed as a frequency distribution in a contingency table. A contingency table lists all categories for the variable and displays the number and/or percentage of individuals observed. An example of a contingency table for vehicle counts is shown in Exhibit A-1. The data can also be displayed graphically, as shown in Exhibit A-2. Continuous data are also summarized in frequency distributions. This is accomplished by grouping the data into meaningful intervals and displaying the number and/or frequency distribution as shown in Exhibit A-3 for speed data. Here again, continuous data are often displayed graphically as shown in the example in Exhibit A-4.

Arithmetic Summations - Another way to describe data is in terms of arithmetic summaries. Qualitative data are often summarized by identifying the category that encompasses the majority of cases, called the mode, and by ranking categories. Referring to Exhibit A-1, the model category is "automobiles," which also ranks first. Quantifiable data are usually summarized in terms of central tendency and variability (dispersion). Central tendency summarizes the center (average or middle

VEHICLE MIX ON AN URBAN ARTERIAL		
Vehicle Type	Number	Percent
Auto	100	50%
Bus	50	25%
Truck	50	25%

Exhibit A-1. Contingency Table for Categorical Data

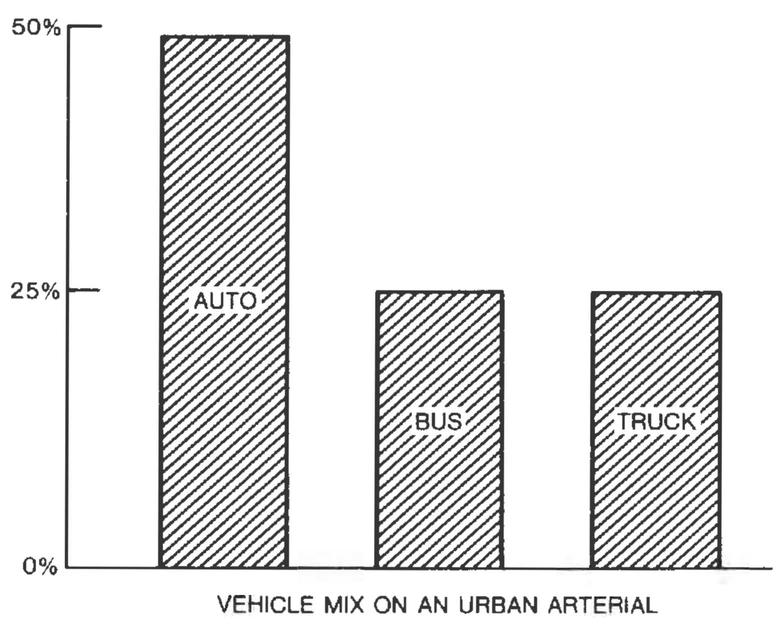
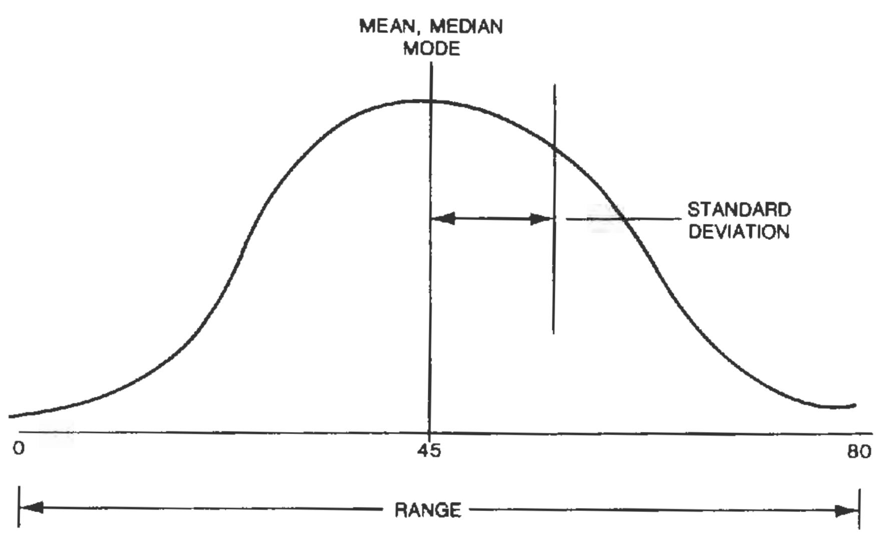


Exhibit A-2.
Categorical Data Displayed Graphically

Spot Speed on a Rural Two-Lane	
Speed (MPH)	Number
0-10	5
10-20	10
20-30	15
30-40	20
40-50	30
50-60	20
60-70	15
70-80	10
> 80	5
TOTAL	150

Exhibit A-3. Frequency Distribution for Continuous Data



SPOT SPEED DISTRIBUTION ON RURAL TWO LANE value) of the data where most observations cluster. It is usually expressed in the mean (average) or median (midpoint). In the data set shown in Exhibit A-3, the mean is determined by adding all speeds and dividing by the number of data points. This yields a mean speed of 45 miles per hour. In the same data set, the median is also 45 miles per hour, obtained by ranking all speeds and finding the midpoint of the middle interval. (In this distribution, called a normal distribution, the mode is also 45 miles per hour.) Variability describes how spread out the data are. Variability is usually expressed by the range or the preferred standard deviation. In many distributions, the mean and median do not coincide. These distributions are referred to as skewed (see Exhibit A-5). There may also be more than one mode (Exhibit A-6). These are called bimodal (two modes) or multimodal (more than two modes) distributions.

Inferential Statistics - The process used to determine if the results are significant is called significance testing. In testing for significance, a supposition (called the null hypothesis) is made concerning the probability of the result occurring by chance. If it can be shown that the probability of the result being a chance occurrence is low (as defined by the significance level chosen), then it can be concluded that the difference does not meet the chance criteria and is, therefore, statistically significant (at the level selected). On the other hand, if there is a higher probability of the results being due to chance than can be tolerated, then the conclusion is made that the difference is not statistically significant. For example, if the average speed before the installation of an advisory sign is 60 miles per hour, and the average speed after the sign is installed is 50 miles per hour, and the chance of a 10-mile per hour difference occurring is one in one thousand (which can be accepted as significant), then the conclusion can be made that the results are probably not due to chance and are, therefore, significant. On the other hand, if there is a one in five chance of obtaining a 10-mile per hour difference (which is not accepted as significant), then it is not possible to rule out the results being due to chance, and the conclusion is made that there is no significance. The probability of the result being due to chance is called the level of significance.

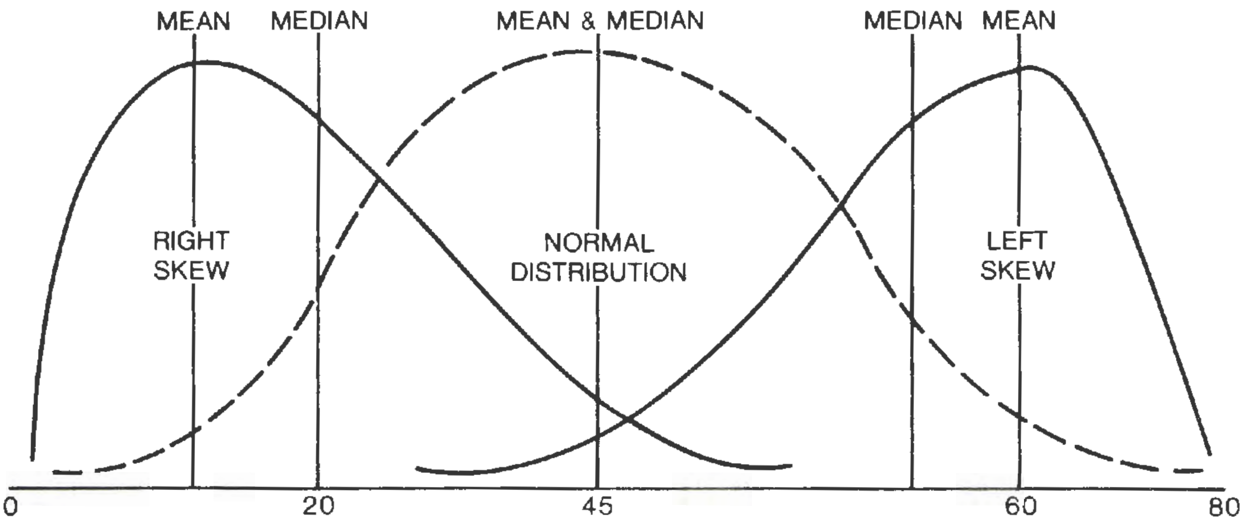


Exhibit A-5.
Examples of Skewed Distributions

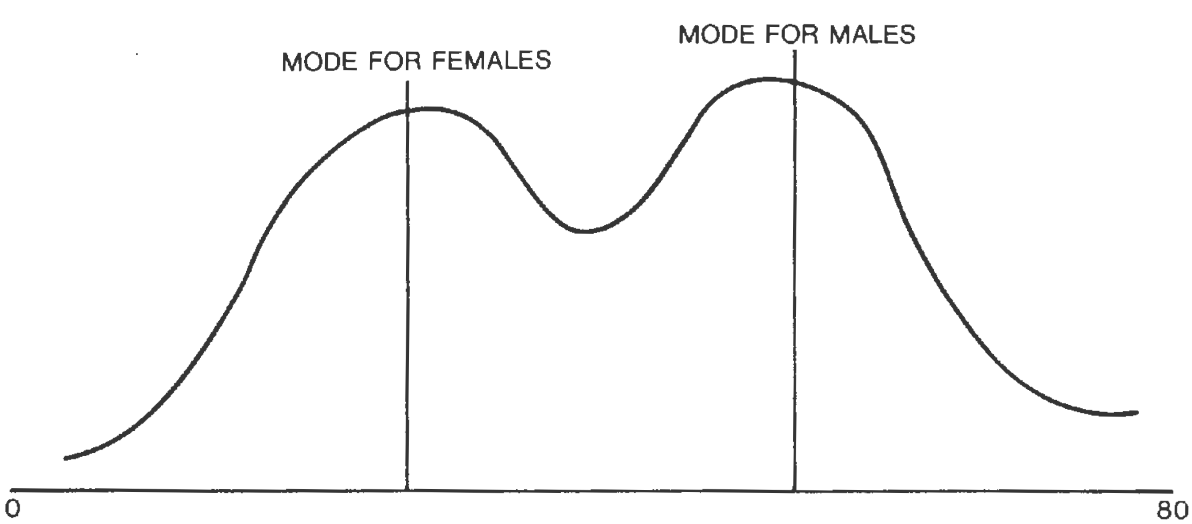


Exhibit A-6.
Example of Bi Modal Distributions

The following are recommended tests of significance:

Poisson Distribution Test -This test is based on the assumption that accidents are rare events with a very low probability of occurrence, distributed in accordance with a distribution called the "Poisson Distribution." The Poisson Distribution test is a test specifically developed to determine if differences in accident rates or frequencies are significant.

t Test - This test is called a parametric test because it is based on the assumption that the MOE sampled during the Data Collection phase approximates the distribution of the overall population (an attribute of a sample is called a variable, an attribute of a population is called a parameter). A further assumption is made that the parameter is distributed in accordance with a distribution called the normal distribution (the normal distribution is "bell-shaped" and looks like the curve shown in Exhibit A-4). Parametric tests require that the data be continuous and interval in nature. Since many of the MOE's (e.g., speed, delay, traveltime) satisfy these assumptions, t is a useful test of significance with wide applicability.

F Test - The F test is also based on parametric assumptions. However, unlike the t test, which tests for significance in means, F is a special test to determine whether differences in variability are significant. For example, there may be little change in average speed before and after an improvement, however, the range of speeds (speed variance) may have been narrowed. F is used to test whether changes in variance, as measured by the standard deviation, are significant.

Z Test - Another specialized parametric test is the Z test. This test is used to determine if differences in large sample (equal to or greater than 30 observations) proportions are significant. Z is used with MOE's such as erratic maneuvers and traffic conflicts.

Chi Square - This test is a nonparametric test because no assumptions have to be met concerning the underlying distribution of MOE's being tested. Because of this, Chi Square can be used to test the significance of any qualitative MOE or any MOE which yields categorical data. Chi Square can also be used in place of the Z test with small sample data (less than 30 observations).

Appendix B - Procedure to Summarize and Compare Accident MOE's

The material in this appendix has been abstracted from a report by Datta, Perkins and Taylor (1978). It is to be used when accident MOE's are evaluated (see Step 19A).

I - PREPARE SUMMARY TABLES FOR ALL DATA

Summary tables are first developed for all data compiled in the Data Collection phase. The Accident Data Comparison form shown in Exhibit B-1 is used to tabulate accident data, accident rates and exposure data. The column headings are modified to indicate the characteristics of the design used, with the "control" columns crossed out for the Before-After design.

When tabulating data for the Control Site design, entries to the "control" columns should be made for each year separately, and a summary table prepared for the entire project evaluation period. Entries should represent the average value of the data for all control sites being considered. When tabulating entries in the data column for the Before-After design, entries should be made for each year separately, and a summary table prepared for the entire project evaluation period.

II - CALCULATE THE PERCENT CHANGE IN THE MOE

An estimate of the expected accident rate or frequency is then made. This estimate is based on the assumptions associated with the design being used. For example, the expected MOE for the Control Site design is based on the accident experience at the control site(s). For the Before-After design, the expected value of the MOE is based on the accident experience at the project site prior to improvement implementation.

The expected value of the MOE can be estimated in two ways, depending on the characteristics of the MOE over time. If the yearly mean values of the MOE follow an increasing or decreasing trend when plotted over several years, the expected MOE should be estimated by using linear regression

ACCIDENT DATA COMPARISON—STEP 19 A

PROJECT: ______

PROJECT NO.: ______

EVALUATOR/DATE: ______

PLAN: ☐ BEFORE/AFTER

☐ CONTROL SITE

Data Summary	Control Before (B_{CF})	Control After (A_{CF})	Project Before (B_{PF})	Project After (A_{PF})	Expected After Rate ____ or Freq. ____	Percent Reduction (%)
Accidents:						
(Fundamental)						
Total Accidents						
Fatal Accidents						
Injury Accidents						
PDO Accidents						
(Project Objectives)						
Exposure						
units: ____ V, or ____ VM						
Comparison Rate ____ or Frequency ____	B_C ____	A_C ____	B_P ____	A_P ____	E ____	(%)
Total Accidents/						
Fatal Accidents/						
Injury Accidents/						
PDO Accidents/						

Exhibit B-1.
Accident Data Comparison Form

techniques. However, if the MOE values follow a horizontal trend or are widely dispersed, the mean value of the MOE over the entire analysis period should form the basis for the expected MOE estimation. The linear regression approach is statistically more attractive; however, its use is subject to: (1) Correlation between the dependent (MOE value) and the independent (time) variables; and (2) the assurance that the slope of the trend line is significantly different from zero (horizontal).

Tests of statistical significance are based on tests of the null hypothesis. This hypothesis states that any difference between two data sets is due to chance. If this can be disproved, then the conclusion is reached that some external factor caused the difference.

For safety project evaluations, the null hypothesis is that the improvement did not affect the accident rate or frequency at the project site, and the accident experience after the improvement is similar to what it would have been if the project were not implemented. The accidents that would have occurred without project implementation cannot be measured (as this condition does not exist) but must be estimated. This estimate is called the expected value and is derived differently for each design. A description of the procedure used to obtain these estimates and the resulting percent change in the MOE are given below.

A - <u>The Control Site Design</u>

1 - <u>Frequency-Related MOE's</u>

a) When the MOE's are frequency-related, and the traffic volumes at the project and/or at the control site are not available, the following equations should be used to compute the expected value of the MOE:

$$E_F = B_{PF} \, (A_{CF}/B_{CF})$$

Where:

E_F = Expected frequency-related MOE at the project site if the improvement had not been implemented.

B_{PF} = "Before" period MOE frequency at the project site.

A_{CF} = "After" period MOE frequency at the control site(s).

B_{CF} = "Before" period MOE frequency at the control site(s).

b) When the MOE's are frequency-related and "Before" volume data are available or can be estimated (see page B-13), the "Before" frequency of accidents, B_{PF} and B_{CF}, must be adjusted for volume changes between the "Before" and "After" period. This is accomplished by multiplying the recorded frequency of accidents in the "Before" period (B_{PF}) by the ratio of AADT "After" to AADT "Before" at the project site. Similarly, at the control site(s), the frequency of "Before" accidents (B_{CF}), would be multiplied by the ratio of the AADT "After" to the AADT "Before" at these sites.

The modified equation for calculating the expected value of the MOE is then:

$$E_F = B_{PF} \frac{(\text{"After" Project AADT})(A_{CF})(\text{"Before" Control AADT})}{(\text{"Before" Project AADT})(B_{CF})(\text{"After" Control AADT})}$$

It is not necessary to adjust the frequencies for dissimilar section lengths between the project site and the control site(s) since the length of section is canceled in using the equations.

c) <u>Percent Change in Frequency Related MOE"s:</u>

Percent change in the frequency related MOE is computed by the following equation:

$$\text{Percent Change} = ((E_F - A_{PF})/E_F)100$$

Where:

E_F = Expected frequency-related MOE at the project site if the improvement had not been implemented.

A_{PF} = "After" period MOE frequency at the project site.

The value for the expected frequency-related MOE, E_F, as described for this and the percent reduction will serve as direct input to the statistical testing procedure (Step 20A).

2 - <u>Rate-Related MOE's:</u>

a) When the MOE's are rate-related and traffic volumes are available or can be estimated (see page B-13) for both project and control sites, the following equations should be used to compute the expected value and percent change in the MOE.

$E_R - B_{PR} (A_{CR}/B_{CR})$

Where:

E_R = Expected rate-related MOE at the project site if the project had not been implemented.

B_{PR} = "Before" period MOE rate at the project site.

A_{CR} = "After" period MOE rate at the control site(s).

B_{CR} = "Before" period MOE rate at the control site(s).

Because the MOE's are expressed in terms of accident rates (as opposed to frequencies), no volume related adjustment is necessary.

b) Percent change in the rate-related MOE is then computed by the following equation.

Percent Change = $((E_R - A_{PR})/E_R)100$

Where:

E_R = Expected rate-related MOE at the project site if the project had not been implemented.

A_{PR} = "After" period MOE rate at the project site.

c) To determine the expected "Before" accident frequency, the expected value of the rate-related MOE must be transformed from an accident rate to an accident frequency. This accident frequency will represent the expected accident frequency. The expected accident frequency is calculated as follows:

$E_F = E_R$ x "After" Project Exposure/10^6

Where:

E_F = Expected "Before" accident frequency to be used in the statistical testing procedure.

E_R = Expected rate-related MOE at the project site if the project had not been made (expressed in accidents/MV or MVM).

Exposure (MV) = Number of vehicles passing an intersection or spot location during the "after" period, expressed in MV (million vehicles).

Exposure (MVM) - Number of vehicles traveling over a section of roadway during the period multiplied by the length of the section, expressed in MVM (million vehicle miles).

3 - Linear Regression Technique

The calculation for the expected value of the MOE as described above is based on the assumption that the value of the MOE is constant over the entire "After" period. If the data for the control site(s) indicates that there may be an increasing or decreasing trend in the MOE over time, a regression technique should be used to determine the expected value (E_F or E_R) of the MOE.

Linear regression is a technique for expressing a linear (straight-line) functional relationship between related variables. Correlation is used to express the precision with which the value of the variable can be predicted if we know the values of the associated variables. The user should be cautioned that just because a functional relationship exists, this does not necessarily mean that a causal relationship exists.

The least square regression technique is recommended for a trend analysis of the MOE. In this technique, the value of the MOE for each year (Y_i) is plotted against time (X_i), where the i represents the number of years from the beginning of the evaluation period. The equation of the line which "best fits" the trend in the MOE is then given by:

$$Y_i = \overline{Y} + b\,(X_i - \overline{X})$$

Where:

Y_i = the estimated value of the MOE in year i.

$\bar{Y}$ = the average value of the MOE over the entire evaluation period.

X_i = the year for which the estimate is desired.

$\bar{X}$ = the mid-point of the evaluation period.

b = the regression coefficient (i.e., slope of the regression line).

The regression coefficient (i.e., slope of the regression line) is obtained from:

$$b = \sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y}) / \sum_{i=1}^{n} (X_i - \bar{X})^2$$

Where:

$(X_i - \bar{X})$ = the value of the difference between each year and the mid-point of the evaluation period (i.e., mid-point of the "Before" plus "After" period).

$(Y_i - \bar{Y})$ = The value of the difference between the MOE for each year and the average value of the MOE over the entire analysis period.

n = the number of years used in the analysis period.

Since the regression technique is designed to test the strength of the relationship between the accident rate or frequency and time, longer time periods yield more reliable results. Therefore, the maximum number of years for which data are available should be used. Further, the maximum number of data points should be used in the analysis. Therefore, it is recommended that the control site MOE for all years ("Before" and "After") and the MOE values for the project site for the "Before" period be used to develop the linear regression model. This will increase the number of data points and ensure that the regression model is representative of the "Before" project site MOE's.

Two tests should be performed to determine whether the indicated trend is significant or is due to random variations in the data. The first test should be an evaluation of the correlation coefficient (r). The square of this coefficient is a measure of the ability of the independent variable (time) to explain the variation in the dependent variable (MOE). If the value of r^2 is greater than 0.8, then use of the regression results should be considered. If r^2 is less than 0.8, then the average value of the control and project site MOE should be used as described previously.

The correlation coefficient can be calculated as:

$$r = S_{xy}/\sqrt{S_{xx}S_{yy}}$$

Where:

$$S_{xx} = n\sum_{i=1}^{n} X_i^2 - (\sum_{i=1}^{n} X_i)^2.$$

$$S_{yy} = n\sum_{i=1}^{n} Y_i^2 - (\sum_{i=1}^{n} Y_i)^2.$$

$$S_{yy} = n\sum_{i=1}^{n} X_iY_i - (\sum_{i=1}^{n} X_i)(\sum_{i=1}^{n} Y_i).$$

Where the variables n, Xi and Yi are as previously defined.

The second test is a determination of the significance of the regression coefficient (b). This test is used to determine whether the slope of the line is significantly different than zero. The equation for this test is:

$$t = (b/S_e)\sqrt{S_{xx}/n}$$

Where: $Se^2 = (S_{xx}S_{yy} - S_{xy}2)/(n(n-2)S_{xx})$

If the value of "t" from this equation exceeds the values in Table B-1, then the regression coefficient (b) is significant, and the regression equation should be used to obtain E_R or E_F, the expected value of the MOE.

TABLE B-1. 't' STATISTIC FOR VARIOUS LEVELS OF CONFIDENCE

Years	"t" Values at Level of Confidence		
n	0.8	0.9	0.95
4	0.941	1.533	2.132
6	0.906	1.440	1.943
8	0.899	1.397	1.860
10	0.879	1.372	1.812
12	0.873	1.356	1.782
14	0.866	1.345	1.761

If the accident trend before project implementation was increasing with time, the use of regression analysis will result in an estimate value higher than that based on the recorded MOE values. It is important that the trend be well established to avoid overestimating project effectiveness. For this reason, it is recommended that the column for a .9 level of confidence be used to enter Table B-1. This requires that we are at least 90% sure that the slope of the trend is different than zero, and thus can be used to estimate future values of the MOE.

The expected value and the percent reduction in the MOE can be calculated using the following equations:

$E_i = \overline{Y} + b(X_i - \overline{X})$

Where:

E_i = Expected MOE at the project site for time period i, if no improvement had been made.

X_i = Years since the beginning of the analysis period.

a) If the MOE's are frequency-related, the equation should be solved for each year of the "After" period and the sum of these MOE's used as the expected MOE frequency for the "After" period.

The percent change is then calculated as follows:

Percent Change = $(E_F - A_{PF})/E_F)/100$

Where:

E_F = Expected frequency related MOE at the project site if no improvement had been made and the expected "Before" accident frequency to be used in the statistical testing procedure.

A_{PF} = The sum of the "After" period MOE frequency at the project site.

b) If the MOE's are rate-related, the equation for E_i should be solved for the midpoint of the "After" period. This value will be the expected MOE rate for the "After" period.

The percent change is then calculated as follows:

Percent Change = $((E_R - A_{PR})/E_R)100$

Where:

E_R = Expected rate-related MOE if the improvement had not been made.

A_{PR} = "After" MOE rate at the project site.

The development of the linear regression analysis may be facilitated by the use of the Linear Regression Summary Table shown in Exhibit B-2.

The expected "Before" accident frequency for statistical testing purposes is calculated as described earlier in this section (page B-5).

LINEAR REGRESSION

PROJECT: ______________________________

PROJECT NO.: ______________________________

EVALUATOR/DATE: ______________________________

Check One: Frequency MOE ________ or Rate MOE ________

X_i Eval. Period (Yrs.) (1)	Y_i Meas. of MOE (2)	$X_i \cdot \bar{X}$ Col. (1) $\cdot \bar{X}$ (3)	$(X_i \cdot \bar{X})^2$ (Col. (3))^2 (4)	$Y_i - \bar{Y}$ Col. (2) $\cdot \bar{Y}$ (5)	$(X_i - \bar{X})(Y_i - \bar{Y})$ Col. Col. (3) × (5) (6)	X_i^2 $\left(\text{Col. (1)}\right)^2$ (7)	Y_i^2 $\left(\text{Col. (2)}\right)^2$ (8)	$X_i Y_i$ Col. Col. (1) × (2) (9)
$\Sigma =$ $\bar{X} =$	$\Sigma =$ $\bar{Y} =$		$\Sigma =$		$\Sigma =$	$\Sigma =$	$\Sigma =$	$\Sigma =$

Exhibit B-2.
Linear Regression Form

B - The Before - After Design

1 - Frequency-Related MOE'S:

a) When the MOE's are frequency-related, and "Before" volume data are available, or can be estimated (see page B-13), the "Before" accident frequency at the project site must be adjusted for volume changes between the "Before" and "After" period. This is accomplished by multiplying the "Before" accident frequency by the ratio of "After" AADT to "Before" AADT.

$$E_F = B_{PF} \text{ ("After" AADT/"Before" AADT)} (T_A/T_B)$$

Where:

E_F = Expected frequency-related MOE at the project site if no improvement had been made.

B_{PF} = The "Before" accident frequency at the project site.

T_A = Length of time of the "After" period.

T_B = Length of time of the "Before" period.

b) In the absence of "Before" volume data, the volume adjustment cannot be made. However, the time period adjustment should be made whenever unequal time periods exist. Thus:

$$E_F = B_{PF}(T_A/T_B)$$

The percent change is then calculated as follows:

$$\text{Percent Change} = ((E_F - A_{PF})/E_F)100$$

Where:

E_F = Expected frequency-related MOE at the project site if no improvement had been made.

A_{PF} = "After" period MOE frequency at the project site.

The value for the expected frequency-related MOE, E_F, will be used directly in the statistical testing procedure as the expected "Before" accident frequency.

2 - Rate-Related MOE's:

a) When the MOE's are rate-related, the expected MOE and percent change is calculated by the following equations:

$$E_R = B_{PR}$$

Where:

E_R = Expected rate-related MOE at the project site if the improvement had not been made.

B_{PR} = "Before" period MOE rate at the project site.

No volume related adjustments are necessary when the MOE's are rate-related.

b) The percent change is then calculated as follows:

$$\text{Percent Change} = (E_R - A_{PR})/E_R)100$$

Where:

E_R = Expected rate-related MOE at the project site if the improvement had not been made.

A_{PR} = "After" period MOE rate at the project site.

c) To determine the expected "Before" accident frequency, the expected rate-related MOE (E_R) must be transformed from an accident rate to an accident frequency. This is accomplished as follows:

$$E_F = E_R \times \text{"After" Project Exposure}/10^6$$

Where:

E_F = Expected before accident frequency to be used in the statistical testing procedure.

E_R = Expected rate-related MOE at the project site if the improvement had not been made (expressed in accidents/MV or MVM).

Exposure (MV) = Number of vehicles passing an intersection or spot location during the period.

Exposure (MVM) = Number of vehicles traveling over a section of roadway during the period multiplied by the length of the section.

If the linear regression technique is used with this plan, the equations for b, r^2 and t are identical to those in the previous design. However, only the data points for the previous "Before" period are used in the regression equations.

d) <u>Estimation of Exposure Index</u>:

There may be times when a project is designated for evaluation after the project has been implemented. When this occurs, accident data are assumed to be available for both the "Before" and "After" period. However, project traffic volumes or exposure data may not have been collected, thereby creating difficulties when rate-related MOE are to be used. This problem may be handled by making an estimate of the before exposure.

The exposure index (MVM or MV) for the period prior to project implementation should be estimated for a point in time equidistant from project implementation to the mid-point to the post-project accident data period. If it is reasonable to assume that the traffic has been increasing or decreasing at a constant rate, then this estimate can be made using:

$$E_b = E_a \left(1/(1+i)^n\right)$$

Where:

E_b = Estimated "Before" period volume (AADT).

E_a = Average volume (AADT) of the "After" period.

i = Average annual traffic growth rate (%).

n = Number of years between the midpoint of the "After" period and the mid-point of the "Before" period.

The average annual traffic growth rate, i, can either be obtained from a knowledge of the growth rate for the city or county in which the project site is located, or it can be estimated using traffic volume data from the "After" period. If data are available from a permanent counter located in the vicinity of the project, the annual growth rate at that station can be used. If no station is located near the site, an estimate of growth rate can be obtained by the following equation:

$$i = (E_2 - E_1)/E_1 \times T_A$$

Where:

E_2 = Traffic volume (AADT) at the end of the "After" period.

E_1 = Traffic volume (AADT) at the beginning of the "After" period.

T_A = Length of the "After" period (in years).

APPENDIX C

Forms

OBJECTIVES LISTING—STEP 1

PROJECT: ____________________

PROJECT NO.: ____________________

EVALUATOR/DATE: ____________________

CATEGORY	OBJECTIVE	COMMENTS
ACCIDENT REDUCTION		
TRAFFIC PERFORMANCE IMPROVEMENT		
SYSTEM PERFORMANCE IMPROVEMENT		

"BEFORE" CONDITIONS —STEP 2

PROJECT: ____________________

PROJECT NO.: ____________________

EVALUATOR/DATE: ____________________

CONDITIONS: WEATHER:

ILLUMINATION:

PAVEMENT:

VOLUME:

SEASON:

OTHERS:

TARGET POPULATIONS:

DRIVER MIX:

VEHICLE MIX:

TARGET DATA COLLECTION PERIODS:

TIME(S) OF DAY:

DAY(S) OF WEEK:

SEASON:

COMMENTS

Page ________ of ________

MOE LISTING—STEP 3

PROJECT: ____________________

PROJECT NO.: ____________________

EVALUATOR/DATE: ____________________

OBJECTIVE:

SITUATION:

POTENTIAL TRAFFIC CONTROL DEVICE(S):

MOE's	MOE TYPE	EXPOSURE UNIT

MOE DEFINITION—STEP 4

PROJECT: ____________________

PROJECT NO.: ____________________

EVALUATOR/DATE: ____________________

MOE:

DATA COLLECTION TECHNIQUE(S):

EQUIPMENT/PERSONNEL REQUIREMENTS:

MOE DEFINITION

OPERATIONAL DEFINITION:

SKETCH:

IMPLEMENTATION AND ACCLIMATION PERIOD—STEP 5

PROJECT: ____

PROJECT NO.: ____

EVALUATOR/DATE: ____

MOE CATEGORIES: ☐ ACCIDENT REDUCTION

☐ TRAFFIC/SYSTEM PERFORMANCE

ESTIMATED START OF IMPLEMENTATION PORTION:

ESTIMATED START OF ACCLIMATION PORTION:

POSSIBLE CHANGES OVER TIME	CONTROLS

TOTAL PERIOD(S):

☐ ACCIDENT MOE's: __________ PLUS 3 YEARS "AFTER" DATA ACCUMULATION

☐ 1 YEAR

☐ OTHER: ______________________________

ESTIMATED END OF TOTAL PERIOD:

EVALUATION DESIGN—STEP 6

PROJECT: ____

PROJECT NO.: ____

EVALUATOR/DATE: ____

DESIGN: ☐ BEFORE-AFTER

☐ BEFORE-AFTER WITH A CONTROL SITE

FOR BEFORE-AFTER WITH A CONTROL SITE DESIGN, IDENTIFY CONTROL SITE(S):

COMMENTS:

STATISTICAL TESTS—STEP 7

PROJECT: ______________________________

PROJECT NO.: ______________________________

EVALUATOR/DATE: ______________________________

MOE: ______________________________

TEST: ______________________________

COMMENTS:

MOE: ______________________________

TEST: ______________________________

COMMENTS:

MOE: ______________________________

TEST: ______________________________

COMMENTS:

CONFIDENCE LEVEL—STEP 8

PROJECT: ____

PROJECT NO.: ____

EVALUATOR/DATE: ____

AGENCY POLICY: ☐ NONE ☐ SPECIFY ________________________________

FACTORS:

PROGRAM IMPORTANCE:	☐ MODERATE	☐ HIGH	☐ VERY HIGH
SAFETY CRITICALITY:	☐ MODERATE	☐ HIGH	☐ VERY HIGH
PROJECT COST:	☐ MODERATE	☐ HIGH	☐ VERY HIGH

CONFIDENCE LEVEL
(LEVEL OF SIGNIFICANCE)

☐ 80% (.20)

☐ 90% (.10)

☐ 95% (.05)

☐ ____________ Other (Specify)

COMMENTS:

SAMPLING PLAN—STEP 9

PROJECT: ____________________

PROJECT NO.: ____________________

EVALUATOR/DATE: ____________________

MOE: ____________________ PERMITTED ERROR ____________________

MINIMUM SAMPLE SIZE: ____________________ DESIRED SAMPLE SIZE: ____________________

SAMPLING PLAN:

MOE: ____________________ PERMITTED ERROR ____________________

MINIMUM SAMPLE SIZE: ____________________ DESIRED SAMPLE SIZE: ____________________

SAMPLING PLAN:

MOE: ____________________ PERMITTED ERROR ____________________

MINIMUM SAMPLE SIZE: ____________________ DESIRED SAMPLE SIZE: ____________________

SAMPLING PLAN:

DATA COLLECTION PLAN AND SCHEDULE—STEP 10

PROJECT: ____________________

PROJECT NO.: ____________________

EVALUATOR/DATE: ____________________

PROJECT LOCATION:

CONTROL SITE LOCATION:

DATE FOR: START OF PROJECT: ____________________

START OF PRE-DATA COLLECTION "BEFORE" PHASE: ____________________

MOE's:

1— ____________________

2— ____________________

3— ____________________

4— ____________________

DATA COLLECTION TECHNIQUE(S):

MOE: ____________________

PRIMARY: ____________________

ALTERNATIVE: ____________________

MOE: ____________________

PRIMARY: ____________________

ALTERNATIVE: ____________________

MOE: ____________________

PRIMARY: ____________________

ALTERNATIVE: ____________________

MOE: ____________________

PRIMARY: ____________________

ALTERNATIVE: ____________________

DATA COLLECTION PLAN AND SCHEDULE	PROJECT NO. ______

EQUIPMENT REQUIREMENTS: (NOTE TYPE, QUANTITY, AVAILABILITY, SERIAL NO.)

MOE: ______

TECHNIQUE: ______

DATES FOR PROCUREMENT ______ OR ASSEMBLY ______

MOE: ______

TECHNIQUE: ______

DATES FOR PROCUREMENT ______ OR ASSEMBLY

MOE: ______

TECHNIQUE: ______

DATES FOR PROCUREMENT ______ OR ASSEMBLY ______

MOE: ______

TECHNIQUE: ______

DATES FOR PROCUREMENT ______ OR ASSEMBLY ______

DATA COLLECTION PLAN AND SCHEDULE	PROJECT NO. ______

CALIBRATION AND COMPARABILITY CHECKS:

EQUIPMENT: ______

SERIAL NO.: ______

REQUIREMENT: ______

EQUIPMENT: ______

SERIAL NO.: ______

REQUIREMENT: ______

EQUIPMENT: ______

SERIAL NO.: ______

REQUIREMENT: ______

EQUIPMENT: ______

SERIAL NO.: ______

REQUIREMENT: ______

DATES FOR CALIBRATION: OR COMBATIBILITY CHECK:

PERSONNEL REQUIREMENTS (NOTE FUNCTIONS, INDIVIDUALS, AVAILABILITY)

DATES FOR RECRUITMENT: OR ASSIGNMENT:

DATA COLLECTION PLAN AND SCHEDULE	PROJECT NO. ______

TRAINING REQUIREMENTS:

DATES FOR TRAINING:

FORM REQUIREMENTS (NOTE FORMS, AVAILABILITY)

DATES FOR DESIGN: ______

TESTING: ______

PRODUCTION: ______

DATES FOR PROCUREMENT: OR ASSEMBLY:

COORDINATION (DESCRIBE REQUIREMENTS)

DATES FOR COORDINATION:

DATA COLLECTION PLAN AND SCHEDULE	PROJECT NO. ______

DATA COLLECTION PROCEDURES, EQUIPMENT LOCATION AND OPERATION, AND PERSONNEL LOCATION AND PROCEDURES, OPERATIONAL DEFINITIONS:

MOE: ______

DATA COLLECTION PLAN AND SCHEDULE	PROJECT NO. ______________

PILOT TESTING CHECK LIST:

(1) NEW OR UNFAMILIAR: EQUIPMENT	☐ YES	☐ NO
DATA COLLECTION TECHNIQUES	☐ YES	☐ NO
PERSONNEL	☐ YES	☐ NO
FORMS	☐ YES	☐ NO
(2) IS TRAINING REQUIRED?	☐ YES	☐ NO
(3) IS DATA COLLECTION COMPLEX?	☐ YES	☐ NO
(4) IS COORDINATION NEEDED?	☐ YES	☐ NO
(5) IS OBTRUSIVENESS A FACTOR?	☐ YES	☐ NO
(6) OTHERS ______________ (SPECIFY)	☐ YES	☐ NO

PILOT TESTING (INDICATE DETAILS)

DATE(S) FOR PILOT TESTING

START ______________

FINISH ______________

DATA COLLECTION PLAN AND SCHEDULES	PROJECT NO. ______

DATE FOR START OF "BEFORE" DATA COLLECTION: ______

DATA COLLECTION CONDITIONS (TIME OF DAY, DAY OF WEEK, ENVIRONMENTAL, ETC.)

CONTINGENCY PLANS:

DATA COLLECTION PLAN AND SCHEDULE | **PROJECT NO.** ____________

SAMPLE SIZE(S):

MOE: ____________

MINIMUM SAMPLE: ____________

MOE: ____________

MINIMUM SAMPLE: ____________

MOE: ____________

MINIMUM SAMPLE: ____________

SAMPLING TECHNIQUE(S): MOE: ____________

MOE: ____________

MOE: ____________

DATA ASSESSMENT (QUANTITY, QUALITY)

DATE(S) FOR DATA ASSESSMENT:

DATE FOR FINISH OF "BEFORE" DATA COLLECTION: ____________

DATA COLLECTION PLAN AND SCHEDULE	PROJECT NO. ______
DATE FOR APPLICATION OF IMPROVEMENT:	
SURVEILLANCE AND CONTROLS DURING IMPLEMENTATION AND ACCLIMATION PERIOD:	
DATE FOR START OF PRE-DATA COLLECTION—"AFTER" PHASE	
PILOT TESTING, TRAINING, ETC. (INDICATE DETAILS)	
DATE(S) FOR PILOT TESTING, TRAINING:	START ______ FINISH ______
DATE FOR START OF "AFTER" DATA COLLECTION:	
DATES FOR DATA ASSESSMENT:	
DATE FOR FINISH OF "AFTER" DATA COLLECTION:	
DATES FOR REDUCTION OF DATA, ASSESSMENT OF RESULTS:	
TARGET DATE FOR FINAL REPORT:	

PRE-DATA COLLECTION ACTIVITIES

PROJECT: ____

PROJECT NO.: ____

EVALUATOR/DATE: ____

PHASE: ☐ "BEFORE" (STEP 11) ☐ "AFTER" (STEP 16)

EQUIPMENT:
- ☐ PROCUREMENT ________________________
- ☐ ASSEMBLY ________________________
- ☐ CALIBRATION ________________________
- ☐ COMPARABILITY CHECK ________________________

PERSONNEL:
- ☐ RECRUITMENT ________________________
- ☐ ASSIGNMENT ________________________
- ☐ TRAINING ________________________

FORMS:
- ☐ PROCUREMENT ________________________
- ☐ ASSEMBLY ________________________
- ☐ DESIGN, TESTING ________________________
- ☐ PRODUCTION ________________________

☐ COORDINATION ________________________

☐ PROCEDURES ________________________

☐ PILOT TESTING ________________________

PILOT TEST RESULTS:

DATA COLLECTION PLAN CHANGES:

DAILY LOG

PROJECT: ____________________

PROJECT NO.: ____________________

EVALUATOR/DATE: ____________________

PHASE: ☐ "BEFORE" (STEP 12) ☐ IMPLEMENTATION AND ACCLIMATION (STEP 15) ☐ "AFTER" (STEP 17)

DATA COLLECTION PERSONNEL:

DATE/TIME	CONDITIONS, EVENTS OR DEVIATIONS

ACCIDENT SUMMARY

PROJECT: ______

PROJECT NO.: ______

EVALUATOR/DATE: ______

DATA SOURCE: ______

TIME PERIOD: ______ TO ______

LOCATION: ☐ PROJECT SITE ☐ CONTROL SITE

☐ "BEFORE"—STEP 13 ☐ "AFTER"—STEP 16

Accident Category	Total Accidents	Fatal Acc.	Fatalities	Injury Acc.	Injuries	PDO Acc.	Invol.
Surface Condition Dry							
Wet							
Snowy Icy							
Other							
Total							
Accident Type Overturn Collision with: Motor Veh.							
Pedestrian							
Pedal Cycle							
Animal							
Fixed Object							
Other							
Total							
Two Veh. Accidents Opposite Direction							
Same Direction							
One Veh. Stopped							
One Veh. Entering Ramp							
One Veh. Exiting Ramp							
Other							
Total							
Two Veh. Accident Types Head-on							
Rear-end							
Sideswipe							
Angle							
Other							
Total							

ACCIDENT SUMMARY

EXPOSURE

Site	Project* Length	Length of Time Period	AADT	Exposure Veh. ____ or Veh. Mi. ____
1.		1.	1.	
TOTAL				

*For vehicle-mile units of exposure (only)

FREQUENCY DISTRIBUTION—CATEGORICAL DATA

PROJECT: ____________________

PROJECT NO.: ____________________

EVALUATOR/DATE: ____________________

MOE: ____________________

PHASE: ☐ "BEFORE"—STEP 13 ☐ "AFTER"—STEP 18

EVENT	FREQUENCY PER ________
TOTAL (Σ)	

ARITHMETIC SUMMARY—UNCLASSED DATA (CONTINUOUS)

PROJECT: ______________________________

PROJECT NO.: ______________________________

EVALUATOR/DATE: ______________________________

MOE: ______________________________

PHASE: ☐ "BEFORE"—STEP 13 ☐ "AFTER"—STEP 18

N	χ	$\bar{X} - \chi$	$(\bar{X} - \chi)^2$	N	χ	$\bar{X} - \chi$	$(\bar{X} - \chi)^2$
TOTALS (Σ)							

$$\bar{X} = \frac{\Sigma\chi}{N} = \frac{\qquad}{} = \underline{\qquad}$$

$$SD = \sqrt{\frac{\Sigma(\bar{X} - \chi)^2}{N}} = \sqrt{\frac{\qquad}{}} = \sqrt{\qquad} = \underline{\qquad}$$

FREQUENCY DISTRIBUTION—ARITHMETIC SUMMARY—CLASSED DATA (CONTINUOUS)

PROJECT: ____________________

PROJECT NO.: ____________________

EVALUATOR/DATE: ____________________

MOE: ____________________

PHASE: ☐ "BEFORE"—STEP 13 ☐ "AFTER"—STEP 18

Class			Class Frequencies (fi)	fiui	fiui²	Relative Frequencies	Cumulative Frequencies	
Boundaries	Limits	Mid-Values (ui)					Number	Relative
TOTALS (Σ)								

FREQUENCY DISTRIBUTION-ARITHMETIC SUMMARY-CLASSED DATA (CONTINUOUS)

MEAN:

$$\bar{x} = \frac{\Sigma \text{ fi } u_1}{\Sigma \text{ fi}}$$

$= \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$

$= \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$

STANDARD DEVIATION:

$$SD = \sqrt{\frac{\Sigma \text{ fi } u_i^2 - \frac{(\Sigma \text{ fi } u_i)^2}{\Sigma \text{ fi}}}{\Sigma \text{ fi} - 1}}$$

$$= \sqrt{\frac{\quad - \frac{(\quad)^2}{\quad}}{\quad - 1}}$$

$$= \sqrt{\frac{\quad}{\quad}}$$

$$= \sqrt{\quad}$$

$= \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$

ACCLIMATION ASSESSMENT—STEP 15

PROJECT: ____

PROJECT NO.: ____

EVALUATOR/DATE: ____

IMPLEMENTATION LENGTH: ____________________________

ACCLIMATION LENGTH: ____________________________

TOTAL LENGTH: ____________________________

POSSIBLE CHANGES OVER TIME	CONTROLS

SURVEILLANCE REQUIREMENTS:

SURVEILLANCE RESULTS:

REQUIRED CHANGES:

MOE COMPARISON—CATEGORICAL DATA—STEP 19

PROJECT: ______________________________

PROJECT NO.: ______________________________

EVALUATOR/DATE: ______________________________

DESIGN: ☐ BEFORE-AFTER ☐ BEFORE-AFTER WITH A CONTROL SITE

MOE: ______________	Project		Control Site	
Event	Before	After	Before	After
	(B_{PR})	(A_{PR})	(B_{CR})	(A_{CR})

NOTE:
FOR BEFORE/AFTER

$E_R = B_{PR}$
$A_R = A_{PR}$

FOR CONTROL SITE
$E_R = B_{PR} \ (A_{CR}/B_{CR})$
$A_R = A_{PR}$

$$\% \text{ CHANGE} = \left[\frac{E_R - A_R}{E_R}\right] \times 100$$

$$= \left[\frac{\quad}{\quad}\right] \times 100$$

$$=$$

ACCIDENT DATA COMPARISON—STEP 19 A

PROJECT: ____________________

PROJECT NO.: ____________________

EVALUATOR/DATE: ____________________

PLAN: ☐ BEFORE/AFTER

☐ CONTROL SITE

Data Summary	Control Before (B_{CF})	Control After (A_{CF})	Project Before (B_{PF})	Project After (A_{PF})	Expected After Rate ____ or Freq. ____	Percent Reduction (%)
Accidents:						
(Fundamental)						
Total Accidents						
Fatal Accidents						
Injury Accidents						
PDO Accidents						
(Project Objectives)						
Exposure						
units: ____ V, or ____ VM						
Comparison Rate ____ or Frequency ____	B_C ____	A_C ____	B_P ____	A_P ____	E ____	(%)
Total Accidents/						
Fatal Accidents/						
Injury Accidents/						
PDO Accidents/						

LINEAR REGRESSION

PROJECT: ____

PROJECT NO.: ____

EVALUATOR/DATE: ____

Check One: Frequency MOE ________ or Rate MOE ________

X_i Eval. Period (Yrs.) (1)	Y_i Meas. of MOE (2)	$X_i \cdot \overline{X}$ Col. (1) $\cdot \overline{X}$ (3)	$(X_i \cdot \overline{X})^2$ (Col. (3))^2 (4)	$Y_i - \overline{Y}$ Col. (2) $\cdot \overline{Y}$ (5)	$(X_i - \overline{X})(Y_i - \overline{Y})$ Col. Col. (3) × (5) (6)	X_i^2 $\left(\text{Col. (1)}\right)^2$ (7)	Y_i^2 $\left(\text{Col. (2)}\right)^2$ (8)	$X_i Y_i$ Col. Col. (1) × (2) (9)
$\Sigma =$ $\overline{X} =$	$\Sigma =$ $\overline{Y} =$		$\Sigma =$		$\Sigma =$	$\Sigma =$	$\Sigma =$	$\Sigma =$

POISSON DISTRIBUTION TEST—STEP 20A

PROJECT: ______________________

PROJECT NO.: ______________________

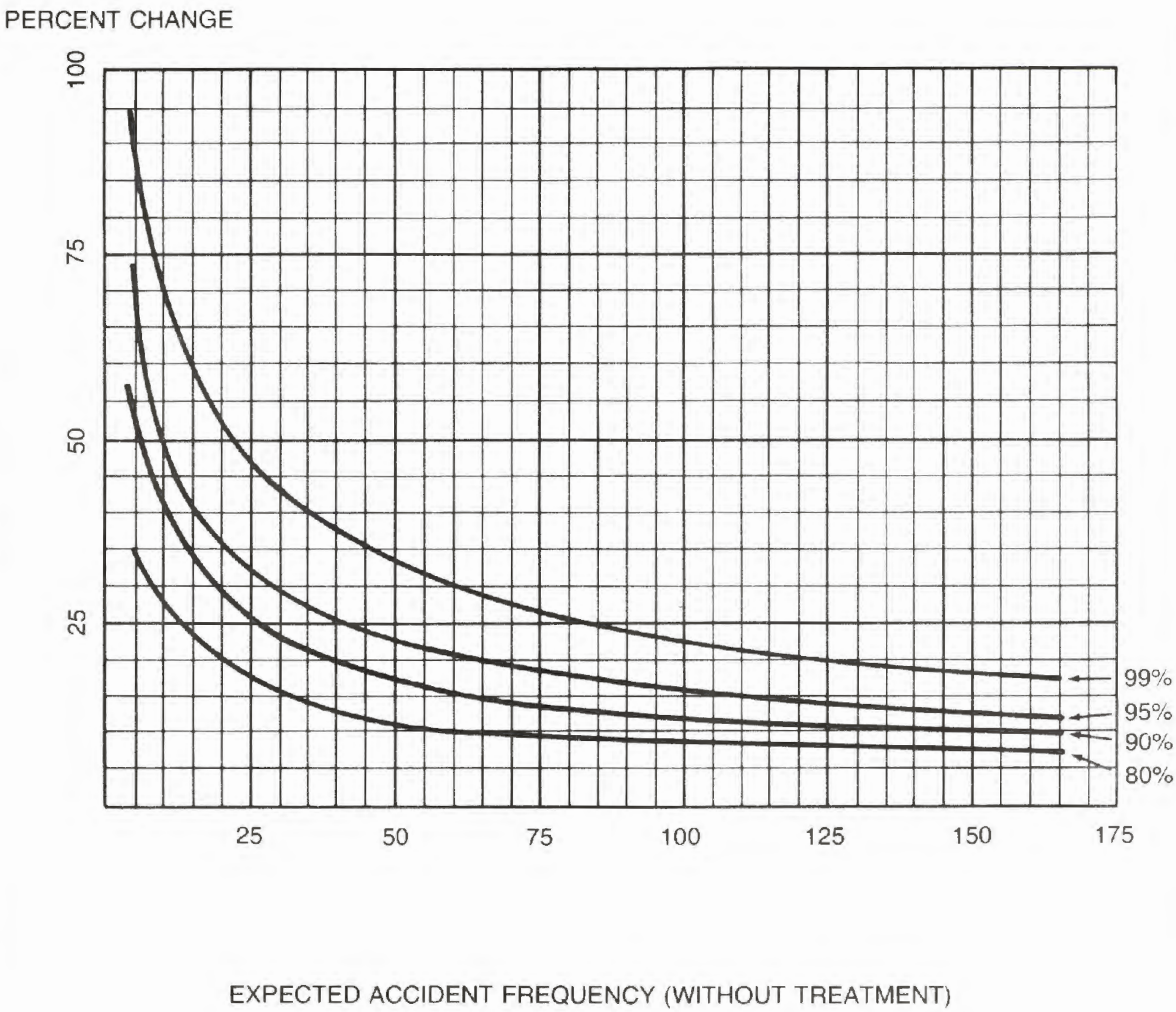
EVALUATOR/DATE: ______________________

CONFIDENCE LEVEL: ☐ 80% ☐ 90% ☐ 95% ☐ Other ____________

Accident Categories	**"After" Frequency ______ Years**		**Percent Reduction**		**Significant For ______ yrs.**
	Observed (A_{PF})	**Expected (E_F)**	**Observed**	**Required**	**Yes or No***
From Step 1	**From Step 19A**		**From Step 19A**	**From Curve At ______ Confidence**	
(Fundamental)					
Total Accidents					
Fatal Accidents					
Injury Accidents					
PDO Accidents					
(Project Objectives)					

*Too small to test

POISSON DISTRIBUTION



t TEST—STEP 20B

PROJECT: ______________________________

PROJECT NO.: ______________________________

EVALUATOR/DATE: ______________________________

MOE: ______________________________

CONFIDENCE LEVEL: ☐ 80% ☐ 90% ☐ 95% ☐ ________ Other

"BEFORE" DATA—FROM STEP 13	"AFTER" DATA—FROM STEP 18
$\bar{X}_B$ = ________	$\bar{X}_A$ = ________
N_B = ________	N_A = ________
SD_B = ________	SD_A = ________

$$t = \frac{\bar{X}_B - \bar{X}_A}{\sqrt{\dfrac{N_B SD_B^2 + N_A SD_A^2}{N_B + N_A - 2}\left(\dfrac{N_B + N_A}{N_B N_A}\right)}} = \frac{\quad - \quad}{\sqrt{\dfrac{(\quad)(\quad)^2 + (\quad)(\quad)^2}{(\quad) + (\quad) - 2}\left(\dfrac{\quad + \quad}{\quad \times \quad}\right)}}$$

$$= \frac{\qquad}{\sqrt{(\quad)(\quad)}} = \frac{\qquad}{\qquad} =$$

DEGREES OF FREEDOM = $N_B + N_A - 2$ = =

t VALUE (FROM TABLE) WITH ________ DEGREES OF FREEDOM

AND ______% CONFIDENCE LEVEL = ____________.

COMPUTED t = ____________

RESULTS: ☐ SIGNIFICANT ☐ NOT SIGNIFICANT

t TEST

VALUES OF t

Degrees of Freedom	Levels of Confidence 80%	90%	95%	99%
1	1.38	3.08	6.31	31.82
2	1.06	1.89	2.92	7.00
3	0.98	1.64	2.35	4.54
4	0.94	1.53	2.13	3.75
5	0.92	1.48	2.02	3.37
6	0.91	1.44	1.94	3.14
7	0.90	1.42	1.90	3.00
8	0.89	1.40	1.86	2.90
9	0.88	1.40	1.83	2.82
10	0.88	1.37	1.81	2.76
11	0.88	1.36	1.80	2.72
12	0.87	1.36	1.78	2.68
13	0.87	1.35	1.77	2.65
14	0.87	1.35	1.76	2.62
15	0.87	1.34	1.75	2.60
16	0.87	1.34	1.75	2.58
17	0.87	1.33	1.74	2.58
18	0.86	1.33	1.73	2.55
19	0.86	1.33	1.73	2.54
20	0.86	1.33	1.73	2.53
21	0.86	1.32	1.72	2.52
22	0.86	1.32	1.72	2.51
23	0.86	1.32	1.71	2.50
24	0.86	1.32	1.71	2.49
25	0.86	1.32	1.71	2.49
26	0.86	1.32	1.71	2.48
27	0.86	1.31	1.70	2.47
28	0.86	1.31	1.70	2.47
29	0.85	1.31	1.70	2.46
30	0.85	1.31	1.70	2.46
40	0.85	1.30	1.68	2.42
50	0.85	1.30	1.67	2.39
60	0.85	1.29	1.66	2.36
∞	0.84	1.28	1.65	2.33

F TEST—STEP 20C

PROJECT: ______________________

PROJECT NO.: ______________________

EVALUATOR/DATE: ______________________

MOE: ______________________

CONFIDENCE LEVEL: ☐ 90% ☐ 95% ☐ ________ Other

"BEFORE" DATA-FROM STEP 13	"AFTER" DATA-FROM STEP 18
SD_B = ______________	SD_A = ______________
N_B = ______________	N_A = ______________

LARGER SD_L = ____________ N_L = ____________

SMALLER SD_S = ____________ N_S = ____________

LARGER VARIANCE = S_L^2 = (LARGER SD)^2 = (   )^2 = ____________

SMALLER VARIANCE = S_S^2 = (SMALLER SD)^2 = (   )^2 = ____________

$$F = \frac{S_L^2}{S_S^2} = \frac{\quad\quad}{\quad\quad} = \_\_\_\_\_\_$$

LARGER DEGREES OF FREEDOM = $N_L - 1$ = ____________

SMALLER DEGREES OF FREEDOM = $N_S - 1$ = ____________

F VALUE (FROM TABLE) WITH ________ DEGREES OF FREEDOM

FROM NUMERATOR (LARGER) AND ________ DEGREES OF FREEDOM

FROM DENOMINATOR (SMALLER) AND ________ % CONFIDENCE LEVEL = ________.

COMPUTED F = ____________

RESULTS: ☐ SIGNIFICANT ☐ NOT SIGNIFICANT

VALUES OF F (.10 LEVEL)

90% Confidence Level

f_1 = DEGREES OF FREEDOM FOR NUMERATOR

f_2 = DEGREES OF FREEDOM FOR DENOMINATOR

f_2	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.20	60.70	61.22	61.74	62.00	62.26	62.53	62.79	63.06	63.33
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48	9.49
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.20	5.18	5.18	5.17	5.16	5.15	5.14	5.13
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.79	3.78	3.76
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14	3.12	3.10
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.90	2.87	2.84	2.82	2.80	2.78	2.76	2.74	2.72
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.67	2.63	2.59	2.58	2.56	2.54	2.51	2.40	2.47
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.34	2.32	2.29
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.38	2.34	2.30	2.28	2.25	2.23	2.21	2.18	2.16
10	3.28	3.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.28	2.24	2.20	2.18	2.16	2.13	2.11	2.08	2.06
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.03	2.00	1.97
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.15	2.10	2.06	2.04	2.01	1.99	1.96	1.93	1.90
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.90	1.88	1.85
14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10	2.05	2.01	1.96	1.94	1.91	1.89	1.86	1.83	1.80
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	2.02	1.97	1.92	1.90	1.87	1.85	1.82	1.79	1.76
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03	1.99	1.94	1.89	1.87	1.84	1.81	1.78	1.75	1.72
17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00	1.96	1.91	1.86	1.84	1.81	1.78	1.75	1.72	1.69
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98	1.93	1.89	1.84	1.81	1.78	1.75	1.72	1.69	1.66
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96	1.91	1.86	1.81	1.79	1.76	1.73	1.70	1.67	1.63
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.89	1.84	1.79	1.77	1.74	1.71	1.68	1.64	1.61
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92	1.88	1.83	1.78	1.75	1.72	1.69	1.66	1.62	1.59
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.64	1.60	1.57
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89	1.84	1.80	1.74	1.72	1.69	1.66	1.62	1.59	1.55
24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.61	1.57	1.53
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87	1.82	1.77	1.72	1.69	1.66	1.63	1.59	1.56	1.52
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.81	1.76	1.71	1.68	1.65	1.61	1.58	1.54	1.50
27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87	1.85	1.80	1.75	1.70	1.67	1.64	1.60	1.57	1.53	1.49
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.79	1.74	1.69	1.66	1.63	1.59	1.56	1.52	1.48
29	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83	1.78	1.73	1.68	1.65	1.62	1.58	1.55	1.51	1.47
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.77	1.72	1.67	1.64	1.61	1.57	1.54	1.50	1.46
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.47	1.42	1.38
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.66	1.60	1.54	1.51	1.48	1.44	1.40	1.35	1.29
120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.60	1.54	1.48	1.45	1.41	1.37	1.32	1.26	1.19
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.55	1.49	1.42	1.38	1.34	1.30	1.24	1.17	1.00

VALUES OF F (.05 LEVEL)

95% Confidence Level

f_1 = DEGREES OF FREEDOM FOR NUMERATOR

f_2 = DEGREES OF FREEDOM FOR DENOMINATOR

f_2	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161.45	199.50	215.71	224.56	230.16	233.99	236.77	238.88	240.54	241.88	243.91	245.95	248.01	249.05	250.09	251.14	252.20	253.25	254.32
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.00	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.48	4.43	4.40	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.84	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.00	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.74	1.79	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.48	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.36	1.23
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.62	1.46	1.39	1.32	1.22	1.00

Z TEST—STEP 20D

PROJECT: ______________________

PROJECT NO.: ______________________

EVALUATOR/DATE: ______________________

MOE: ______________________

CONFIDENCE LEVEL: ☐ 80% ☐ 90% ☐ 95% ☐ ____________ Other

n_1 = "Before" erratic maneuvers or traffic conflicts = ____________

n_2 = "After" erratic maneuvers or traffic conflicts = ____________

N_1 = Traffic Volume – "Before" = ____________

N_2 = Traffic Volume – "After" = ____________

$n = n_1 + n_2$ = ____________

$N = N_1 + N_2$ = ____________

$$Z = \frac{N_2 n_1 - N_2 n_1}{\sqrt{\dfrac{N_1 N_2\, n\,(N - n)}{N}}} = \frac{\qquad}{\sqrt{\dfrac{\qquad(\qquad)}{\qquad}}}$$

=

VALUES OF Z	CONFIDENCE LEVEL			
	80% = 0.842	90% = 1.262	95% = 1.645	99% = 1.960

Z VALUE (FROM TABLE) WITH ____________%

CONFIDENCE LEVEL = ____________

COMPUTED Z = ____________

RESULTS: ☐ SIGNIFICANT ☐ NOT SIGNIFICANT

CHI SQUARE—STEP 20E

PROJECT: ____________________

PROJECT NO.: ____________________

EVALUATOR/DATE: ____________________

MOE: ____________________

Event	"Before" fo	"After" fh	fo−fh	$(fo-fh)^2$	$\frac{(fo-fh)^2}{fh}$
TOTAL (Σ)					

DEGREES OF FREEDOM = NUMBER OF EVENTS −1 = ____________

CONFIDENCE LEVEL: ☐ 80% (.20) ☐ 90% (.10) ☐ 95% (.05) ☐ ____________ Other

CHI SQUARE VALUE (FROM TABLE) WITH ____________ DEGREES OF FREEDOM

AND ____________ % CONFIDENCE LEVEL = ____________.

COMPUTED CHI SQUARE = ____________

RESULTS: ☐ SIGNIFICANT ☐ NOT SIGNIFICANT

CHI SQUARE

CHI SQUARE

df	Levels of Confidence 80%	90%	95%	99%
1	1.64	2.71	3.84	6.64
2	3.22	4.61	5.99	9.21
3	4.64	6.25	7.82	11.34
4	5.99	7.78	9.49	13.28
5	7.29	9.24	11.07	15.09
6	8.56	10.65	12.59	16.81
7	9.80	12.02	14.07	18.48
8	11.03	13.36	15.51	20.09
9	12.24	14.68	16.92	21.67
10	13.44	15.99	18.31	23.21
11	14.63	17.26	19.68	24.72
12	15.81	18.55	21.03	26.22
13	16.99	19.81	22.36	27.69
14	18.15	21.06	23.68	29.14
15	19.31	22.31	25.00	30.58
16	20.47	23.54	26.30	32.00
17	21.62	24.77	27.59	33.41
18	22.76	25.99	28.87	34.80
19	23.90	27.20	30.14	36.19
20	25.04	28.41	31.41	37.57
21	26.17	29.62	32.67	38.93
22	27.30	30.81	33.92	40.29
23	28.43	32.01	35.17	41.64
24	29.55	33.20	36.42	42.98
25	30.68	34.38	37.65	44.31
26	31.80	35.56	38.88	45.64
27	32.91	36.74	40.11	46.96
28	34.03	37.92	41.34	48.28
29	35.14	39.09	42.56	49.59
30	36.26	40.26	43.77	50.89

☆U.S. GOVERNMENT PRINTING OFFICE: 1980-628-786