UM-HSRI-SA-73-9-1&2

EVALUATION OF FORD EMPLOYE SKILLED DRIVING PROGRAM

Volumes 1&2

Summary and Technical Report

Martin E. Lee

HIGHWAY SAFETY RESEARCH INSTITUTE
UNIVERSITY OF MICHIGAN

October 1973 - February 1974

Under contract to: FORD MOTOR COMPANY

BIBLIOGRAPHIC DATA SHEET	1. Report No. UM-HSRI-SA-73-9	2.	3. Recipient's Accession No.

4. Title and Subtitle
Evaluation of Ford Employe Skilled Driving Program. Volumes 1 & 2. Summary and Technical Report.

5. Report Date
Oct. 1973-Feb. 1974

6.

7. Author(s)
Martin E. Lee

8. Performing Organization Rept. No. UM-HSRI-SA-73-9

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

10. Project/Task/Work Unit No.

11. Contract/Grant No.

12. Sponsoring Organization Name and Address
Ford Motor Company
Dearborn, Michigan 48121

13. Type of Report & Period Covered Final Report
Feb. 1972-Feb. 1974

14.

15. Supplementary Notes

16. Abstracts This project is an independent evaluation of a driver improvement program for unexceptional adult drivers, offered on a pilot basis in 1972-73 by Ford Motor Company to its employes. The program was conducted using four treatment groups which received different amounts of re-training (N=750). This report discusses the development and application of an evaluation model employing immediate criteria (operational success and survey results), intermediate criteria (performance tests before and after training), and ultimate criteria (crash and loss rates). Immediate and intermediate data are discussed together with comparisons of the treatment groups on biographical and driving record characteristics. It was found that two treatments consisting of several hours of classroom and on-street instruction resulted in 30%-40% increases in mean group score on an observational test of on-street performance. A smaller but still statistically significant increase was found for a third treatment comprising a self-teach course, but a control group showed no change. Small mean increases in headway were recorded for the three treated groups with marginal statistical support. The remaining tests revealed few significant differences. Limited re-testing three months after post-test suggested that the measured improvements were sustained for the two groups who were actively trained.

17. Key Words and Document Analysis. 17a. Descriptors

17b. Identifiers/Open-Ended Terms

17c. COSATI Field/Group

18. Availability Statement	19. Security Class (This Report)	21. No. of Pages
UNLIMITED	UNCLASSIFIED	183
	20. Security Class (This Page) UNCLASSIFIED	22. Price

UM-HSRI-SA-73-9-1

EVALUATION OF FORD EMPLOYE SKILLED DRIVING PROGRAM

VOLUME 1

Summary

Martin E. Lee

HIGHWAY SAFETY RESEARCH INSTITUTE
UNIVERSITY OF MICHIGAN

February 1974

Under contract to: FORD MOTOR COMPANY

FOREWORD

by Ford Motor Company

For two decades, Ford Motor Company has conducted and supported driver behavior research and has developed driver training materials for beginning and experienced drivers. Efforts to quantify elements of driver performance have led to the development of vehicle-installed instrumentation to study drivers. Against this background, early in 1971, Ford launched a multiphase program designed to improve driver licensing procedures, which appear to be a key factor in reducing traffic accidents. This Driver Improvement and Licensing Program was conceived by Ford's Traffic Safety Programs Department as requiring both research and development.

Phase I, called the Ford Employe Skilled Driving Program was designed as a research project to test four basic hypotheses, using company employes as participants.

- The performance of experienced drivers can be improved through training and practice.
- Driver performance can be measured.
- Measurable improvement can be accomplished quickly and economically.
- Improved driver performance will reduce accident probability.

Later phases will test these hypotheses in the field, using both fleet drivers and participants from the general driving population.

To gain a higher level of objectivity and the advantage of specialized expertise for the Phase I research project, the Traffic Safety Programs Department negotiated service contracts with the Highway Safety Research Institute (HSRI) at The University of Michigan and with the American Academy of Transportation (AAT) of Ann Arbor, Michigan.

Under its contract, HSRI agreed to perform the following:

I. Define the experimental plan.

II. Coordinate with and advise Ford project manager throughout the program.

III. Prepare a procedural manual covering items I and II above.

IV. Receive data from the Ford project manager and AAT, obtain driver data from the Michigan Secretary of State, construct a data file, and analyze the data, subjecting them to all practicable statistical techniques.

V. Write a final report, and make oral presentations as required.

AAT contracted to reduce paper tape recorded data from vehicle installed driver performance instrumentation and to provide the services of six persons who could be trained to serve as instructors under the direction of Phil Gram, Ford project manager.

This is the final report prepared by HSRI on this driver research project.

Fletcher N. Platt
Traffic Safety Programs Manager

ACKNOWLEDGMENTS

This evaluation was possible only through the assistance of a sizable group of people.

While at HSRI, Dr. William L. Carlson, now Associate Professor, St. Olaf College, Minnesota, was jointly responsible with the author for the experimental design and made the initial recommendations for analysis procedures. James O'Day, Head, Systems Analysis, HSRI, was Project Director, and most helpful throughout the evaluation. Dr. Richard Kaplan, Research Psychologist, HSRI, and Dr. Donald Smith, Michigan State University, assisted in test development. Dr. Jairus Flora, Assistant Professor of Biostatistics, consulted on statistical techniques, and carried out several of the analyses on the Perception of Hazards and Unusual Uses tests, as well as on the follow-up periods for crash data. Dr. Samuel Schultz II, Associate Professor of Psychology in Nursing was also a valued consultant on statistical techniques. Dr. Arthur C. Wolfe, Senior Research Associate, HSRI, advised on survey design. Mr. John Mahone of HSRI built the data files and carried out many of the computer analyses.

To Fletcher N. Platt, Manager, Traffic Safety Programs, Ford Motor Company; to Mr. Phil Gram, Manager of the Program (now Driver Safety Programs Manager) and his staff, and to the instructors, the author is especially indebted for their painstaking responses to our requests to execute many difficult procedures necessary to maintain the integrity of the experimental design.

Several other people at Ford Motor Company and the American Road Insurance Company were very generous with their time in facilitating our use of the various types of insurance data. These included Mr. Grant Wood, Mr. Anthony Matzdorf, and Mr. Robert Ingersoll.

CONTENTS

INTRODUCTION

Until recently, most efforts to improve or retrain drivers have concentrated upon those identified as "deviant" or "negligent." The rationale was developed that, if sufficient accident involved drivers could be retrained (or prevented from driving), the number of accidents owing to human error could be markedly reduced. However, while the value of rehabilitating exceptionally "poor" drivers is beyond question, it is becoming increasingly clear that this will have little impact on the overall traffic accident problem. In a California study, Burg (1970)[1] suggests that removal of all drivers with two or more accidents in a three year period would eliminate 3.9% of the drivers and 8% of the accidents in the ensuing three years.

Retraining programs aimed at the vast majority of "unexceptional" drivers would therefore appear to be worth attempting, and have interesting implications for improving the driver licensing process. Historically, this approach has not been popular, however, partly because of the greater political viability of the "deviant driver" approach, and partly because of the lack of information on how to efficiently approach driver retraining on such a large potential scale. Although driver programs for company fleets yield some information on large groups of drivers, they are generally atypical of the general population. These also operate in conjunction with sanctions which are rarely viable for the general public and which at best obscure the effects of training programs.

Against this setting, Ford Motor Company launched its multi-phase program of research and development in driver improvement and licensing. The purpose of the Phase I research project was to examine the effectiveness of a set of training procedures designed for unexceptional drivers and, implicitly,

to test the feasibility of offering such training as a program for the employees of a large corporation. This required that considerable effort be addressed to an experimental design format for the program and to the problems of driver performance measurement.

DESCRIPTION OF THE FORD EMPLOYE SKILLED DRIVING PROGRAM

EXPERIMENTAL FORMAT

In developing improvement procedures for unexceptional motorists, it is of critical importance to discover not only whether retraining has any effect on driving behavior, but also how much effect is brought about by different amounts of retraining. Rather large numbers of participants are desirable to investigate several intensities of retraining; moreover, they should be drivers from the widest possible ranges of age and socio-economic background. Because of practical limitations, the program was confined to salaried employes; however, the full range of job grades from junior clerical staff to senior executives was well represented. The program was designed to accommodate 1,052 employes who accepted an invitation to participate in an innovative advanced driving program within the company. This invitation was mailed to 2,000 Detroit area employes who were randomly selected by computer in roughly equal numbers from the managerial and the general salary rolls. Ford has approximately 10,000 salaried employes in the metropolitan area, of which about 2,500 are in the managerial grades.

Figure 1 summarizes the experimental design. The volunteers were randomly assigned to four treatment groups, using methods which ensured that the age and sex distributions were similar from group to group. The treatments comprised three intensities of training and a control. Because of the amount of contact between all participants, the control group could

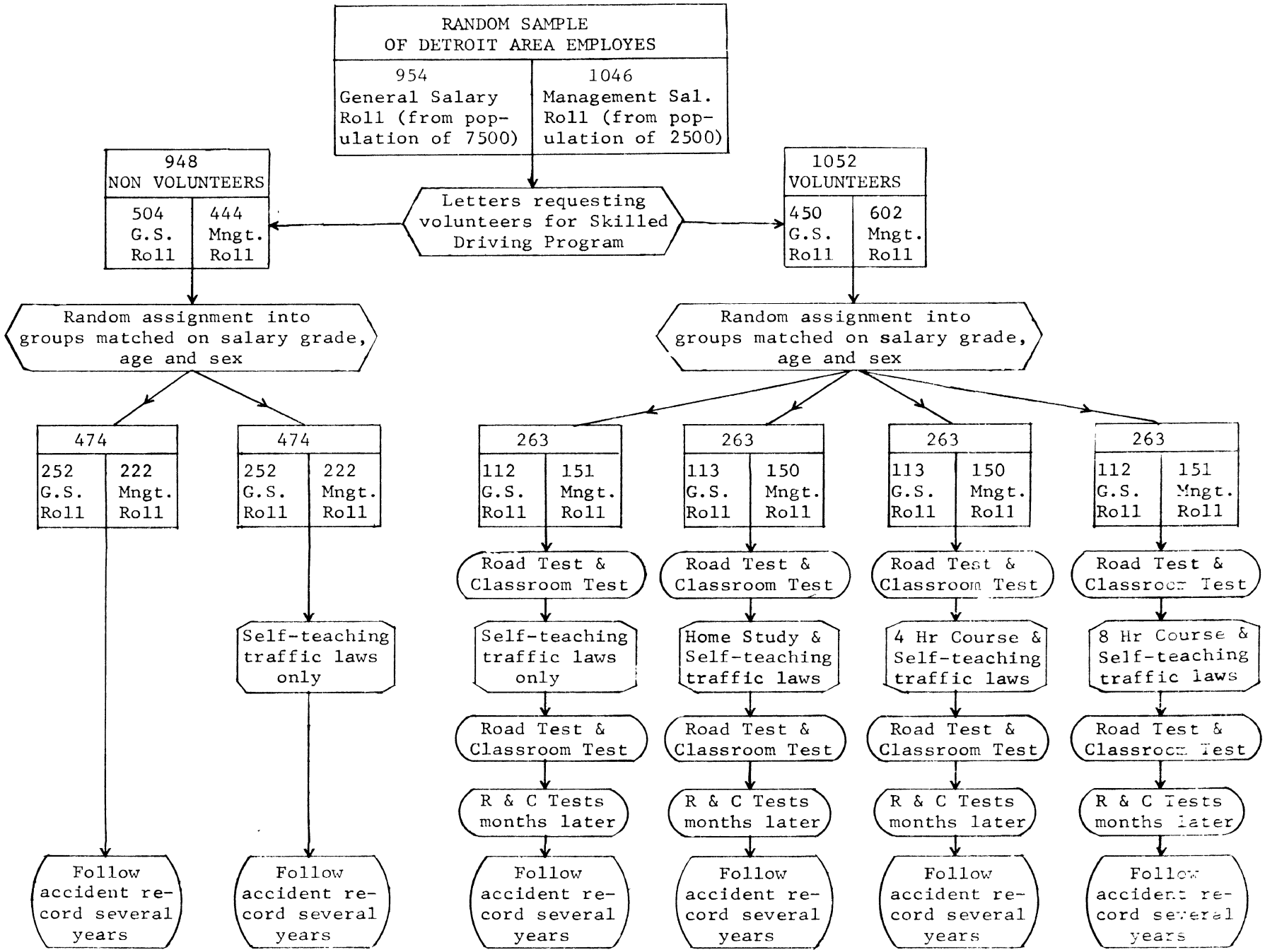


FIGURE 1 The Ford Employe Skilled Driving Program Experimental Format

not be isolated from the knowledge disseminated in the program. This group was therefore given a copy of the new edition of "What Every Driver Must Know," Michigan's state driver manual (as were all participants) and a letter exhorting them to study it and to test themselves using a supplied multiple-choice quiz; also, all four groups were given access to an audio-visual library containing considerably more material than that used in the training sessions. Differences in learning measured between groups could thus reasonably be attributed to factors other than the general arousal of interest in driving information, or the novelty of involvement in the program. The three trained groups were given, respectively, a home study course, a four-hour course, and an eight-hour course. The latter two treatments were made up of two-hour training sessions equally divided between classroom and behind-the-wheel instruction.

All volunteer participants were treated alike with respect to testing. A package of classroom and behind-the-wheel tests was administered as a pre-test; immediately following, participants were oriented to the treatment to which they were assigned. Five weeks later, usually at the same time of day and on the same day of the week, a similar test package was administered (with some additions) as a post-test. A random selection of about 20% of those post-tested were tested again three months after post-test. Arrangements were made to examine previous driving record and to follow post-course record for several years.

Although it was not possible in this phase to train any of the non-volunteers, a sample of driving records are being followed for comparison with the volunteer groups. Moreover, to provide an additional treatment analogous to a no-threat "warning letter," a randomly selected half of the non-volunteers were mailed the same materials as received by the volunteer control group. However, none of the non-volunteers underwent testing or other personal involvement with the program.

EVALUATION MODEL AND TESTING TECHNIQUES

Hitherto, most driver programs have been evaluated on a before and after comparison of driving records. Attractive as this may be as the ultimate criterion of success, its use has proved very problematical. In particular, accidents are sufficiently rare events that very large treatment group sizes are necessary to make statistically valid comparisons of even large changes in rates of occurrence over as little as two or three years. Furthermore, driver record data are notoriously subject to reporting biases. Nevertheless, neither is it reasonable to evaluate a driver program, as is sometimes done, by comparing operational data, such as the number of attendees, acceptability to students, and the like. Such information is of interest and was gathered for this study using program records and a mailed survey; however, it says little or nothing about the changes in driver behavior brought about by the program. It is therefore necessary to use what is generally called "intermediate" measurement to link the application of a training program to any change in accidents or other ultimate desired outcomes.

For this purpose, a "causal chain" model developed earlier at HSRI (O'Day, et al., 1971)[2] was extended to meet the needs of this evaluation.* Figure 2 summarizes this model, which defines a series of stages, corresponding to increasingly complex levels of behavior at which change should be detectable if the program is effective. Figure 2 also shows conceptually that the likelihood that factors other than the program explain measured changes, increases with the remoteness of test criteria from the observable effects of the program. Recognizing some considerable difficulties owing to the poor state-of-the-art of driver testing, the following criterion tests were defined for each stage of change:

---

*A more detailed rationale for this approach than is given in the Technical Report (Volume 2) will be found in Lee (1973).[3]

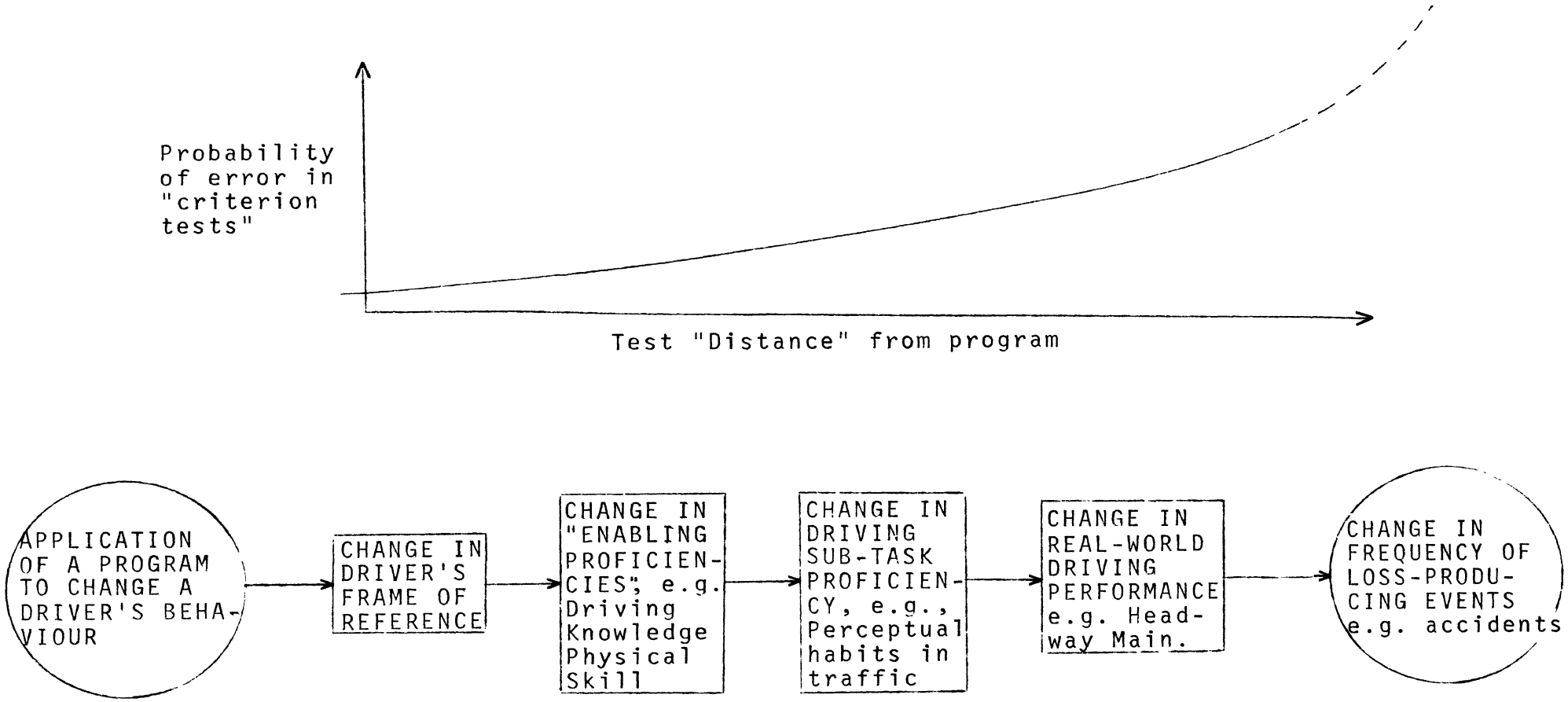


FIGURE 2

A DEVELOPMENT OF THE CAUSAL CHAIN
MODEL FOR DRIVER PROGRAM EVALUATION

INTERMEDIATE EVALUATION	
Change in driver's frame of reference	Multiple-choice test on information emphasized in course. (Written test, Parts I & IV)
Change in "enabling proficiencies"	Multiple-choice test on areas of general driving knowledge related to course content. (Written test, Parts II & III)
Change in "driving subtask proficiencies"	Test of ability to detect hazards in slides of highway traffic situations shown for only five seconds. ("Perception of Hazards")
	Test of imaginative thinking about potentially hazardous driving situations. ("Unusual Uses")
Change in real-world driving performance	Objective (electro-mechanical) measurement of driver actions over a standard route--total elapsed time, steering reversal rates, brake applications and combinations of these. ("Objective Behind-the-Wheel Test")
	Subjective rating of on-street behaviors related to course content, including headway in seconds. ("Subjective Behind-the-Wheel Test")
ULTIMATE EVALUATION	
Change in frequency of loss producing events	Secretary of State driving record. Insurance claims records for those senior employes entitled to a leased car as a job benefit.

COURSE CONTENT AND TRAINING PROCEDURES

The three intensities of training utilized the same basic content. This was developed primarily from the more advanced training materials produced by Ford in recent years with the assistance of the company's driver education consultants.*

*In order to maintain an independent evaluation, HSRI was not involved in the development of the instructional program.

The main theoretical bias implicit in the selection of content was that adults are served better by efforts to improve their ability to think systematically about essential information while driving, than by attempts to drill them in specific procedures for traffic maneuvers. Central to the teaching included in this program was a paradigm for analyzing traffic situations known as the "Decision Pattern." This was closely tied to other selected principles which may be generalized to most normal driving conditions; these include headway timing, maintenance of a space cushion, and associated visual habits. Specific driving maneuvers and highway traffic situations were included to exemplify the application of these principles. In addition, a moderate amount of emphasis was given to emergency and abnormal driving conditions; and general safety habits, such as belt usage, were encouraged.

Testing and training sessions took place on company time at a specially built facility in the lobby of the Ford World Headquarters office building in Dearborn. The audio-visual library was included in this facility. Six heterogeneous instructors were selected to encourage a range of teaching styles, and were trained, as a group, to operate the program. Three of these did not remain throughout the project; two replacements were hired and given on-the-job training.

SUMMARY OF FINDINGS

OPERATIONAL SUCCESS

In our opinion, this program has demonstrated that it is feasible to operate fairly large driver programs in a company setting under experimental controls essential for evaluation. The major operational considerations were:

Maintenance of Sampling

The original response of 1,052 volunteers was high,

yielding 53% of a random sample comprising approximately 20% of all salaried employes in the Detroit area. However, attrition from this group was substantial, with 750 employes attending at least a pre-test, and only 560 participating through the post-test. Higher participation levels are much to be desired, but in this case it could not be shown from an analysis of bio-graphical and driving record characteristics that the volunteers who took part were significantly different, as a group, from those who cancelled or dropped out. More importantly, the sample trained could not, for practical reasons, include any non-volun-teers; this tends to limit the generalizability of the results. Comparisons on age, salary grade, race, sex, marital status and years of service with company, suggested that those who took part did not markedly differ, as a group, from Ford salaried employes as a whole. However, it is not known the extent to which the participant sample, with its tendency towards the middle salary grade and age ranges, can represent the general driving public.

Training Procedures

As noted, differences in teaching style were encouraged. Hence, the two groups which underwent classroom and in-car instruction were vulnerable to bias from disparities in the allocation of instructors. Equitable distribution was made more difficult by the personnel changes. Analyses of class records showed that instructors were close to randomly dis-tributed over the in-car sessions, but not the classroom sessions. Because of greater standardization in content, and lower demands on the instructors' interpersonal skills, dis-parities of instructor distribution are probably less critical to the classroom sessions.

While the intention was not to evaluate in detail the training situations, it should be noted that there was a great deal of variability over time in the quality of the training procedures.

Testing Procedures

Much was learned in the program about the application of tests to this kind of evaluation. Cognitive tests are probably not a problem unless program content becomes much more technical in nature. The classroom tests in the perception and decision-making areas suffered a variety of problems, including some mechanical difficulties. Better measures in these areas are essential if full benefits are to be obtained from causal chain methodology. Their inclusion was worthwhile, however, for the information they yield on the response of different subgroups of adults to the training and testing situations.

The objective in-car measures had not heretofore been applied on a large scale under experimental constraints; such applications are essential for their development. The present data are difficult to interpret in the context of the evaluation design, but as intended, considerable potential exists for a posteriori analyses. The subjective measures were appropriate for this program, but they need to be made less ambiguous for inexperienced observers. Some of these, notably headway, could perhaps be augmented with parallel objective devices, even if they were used only to improve the training of observers. The use of instructors as observers within the same program is definitely to be discouraged.

TEST AND SURVEY RESULTS

Test Results

The major findings are derived from analysis of pre-test and post-test data. Follow-up (three month) test results revealed no substantial reversals of these trends; however, it should be noted that in some follow-up analyses treatment group sizes were very small.

The smallness of differences in both the Content Acquisition and Driving Knowledge portions of the Written Test suggests that the transfer of key information from programs such as this may

well be achieved with little instructional effort. However, all groups showed unexpected deficiencies in general driving knowledge, even though all participants were given a copy of the new Michigan driver's manual.

The Perception of Hazards test was unable to detect group differences in performance resulting from the course. However, the Unusual Uses technique showed small improvements in the ability of the Eight-Hour and Four-Hour participants to think imaginatively about driving.

The most favorable results were provided by the subjective rating of behind-the-wheel performance. Within this test, the mean sum scores for the performance of behaviors logically related to accident-free driving increased between pre- and post-test 40%, 30%, and 10% respectively, for the Eight-Hour, Four-Hour, and Self-Teach groups, compared to a slight decrease for the control group. The three treated groups appeared to increase car following time (headway) by between one-fifth and one-quarter of a second. The sum scores were substantiated using several statistical techniques; the improvements in car following time (headway) were marginally supported by similar analyses. However, it is probable that part, but not all, of the differences recorded are explainable by observer bias, as it proved impossible to conceal the treatment group of subjects from the observers. The objective measures offered only a few weak indications of treatment group differences.

In general, although the test results improved with the intensity of training, the greatest increase occurred between the Self-Teach and Four-Hour groups, that is between those who were not formally trained and those who were. Support exists in the data for conducting several hours of training, but it can not be shown that eight hours were substantially better than four.

Survey Results

The survey data reveals generally high approval of the

concept of advanced driver training for adults and a range of preferences as to how and when the respondents would prefer to undergo it. There was considerable endorsement of behind-the-wheel training and of audio-visual instructional materials, but the overall demand was for some extensions in content and for a variety of learning situations.

RECOMMENDATIONS

1. The Case for Rigorous Evaluation

Because of the great number of factors which influence driver performance, even small amounts of change resulting from improvement and retraining programs are meaningful if they can be supported by rigorously controlled evaluation. Indeed, dramatic amounts of change have not been proven in the literature for large numbers of drivers. We would contend that an evaluation of this complexity is essential to provide accurate information on the amounts and types of training which can be supported as efficient on a large scale. We recommend a continuing effort to assess the effectiveness of methodologies as they are developed, and before they are implemented.

2. The Importance of the Accident Data

Although the sample sizes are small, and therefore very substantial improvements in accident and loss rates must be achieved to be supportable statistically, the driver record follow-up should be continued at least through July 1975. There is no reason to suspect that such developments as the energy shortage and no-fault insurance would affect treatment groups differentially.

3. The Data Base should be Further Utilized

The HSRI data base on the Employ Skilled Driving Program is a valuable resource for further study of adults' responses to

such offerings. It is a product of procedures more rigorous than are generally applied to programs involving large numbers of "average" adult drivers.

4. *The Next Step in this Activity should be to Attempt to Match Unexceptional Drivers to a Variety of Learning Situations Using Limited Areas of Content*

HSRI's key recommendation for future activity in the area of adult driver retraining is that, given a reasonable definition of appropriate instructional content, much attention should be directed to identifying "target groups" of drivers for a variety of instructional procedures. For large populations of unexceptional drivers, training programs must be limited to a few hours; under these circumstances, and given the state-of-the-art of testing, it is unrealistic to expect to develop truly individualized instruction for all drivers in the near future. However, with due attention to the acceptability of various training approaches, including those used in this program, and to the compliance to driving standards which is implicit in different training situations, the characteristics of such target groups should begin to emerge. At that point, it becomes feasible to contribute a great deal towards the ideal setting for the periodic upgrading of drivers -- the process of driver licensing.

REFERENCES

[1]Albert Burg, "The Stability of Driving Record Over Time," Accident Analysis and Prevention, Volume 2, May 1970, Pergamon Press, London, 1972. pp. 57-65.

[2]J. O'Day, J. S. Creswell, Jr., J. A. Green, M. E. Lee, J. C. Marsh, IV, and S. Schultz II. "Planning for the Concentrated Implementation of Highway Safety Countermeasures." Highway Safety Research Institute, University of Michigan, August 1971. 4 Volumes.

[3]Martin E. Lee, "Detecting the Effectiveness of Driver Retraining Among Detroit Car Commuters," Universities Transport Study Group Conference. Proceedings. University College, London. January 3-4, 1973. January 1973.

UM-HSRI-SA-73-9-2

EVALUATION OF FORD EMPLOYE SKILLED DRIVING PROGRAM

VOLUME 2

Technical Report

Martin E. Lee

HIGHWAY SAFETY RESEARCH INSTITUTE
UNIVERSITY OF MICHIGAN

October 1973

Under contract to: FORD MOTOR COMPANY

FOREWORD

by Ford Motor Company

For two decades, Ford Motor Company has conducted and supported driver behavior research and has developed driver training materials for beginning and experienced drivers. Efforts to quantify elements of driver performance have led to the development of vehicle-installed instrumentation to study drivers. Against this background, early in 1971, Ford launched a multiphase program designed to improve driver licensing procedures, which appear to be a key factor in reducing traffic accidents. This Driver Improvement and Licensing Program was conceived by Ford's Traffic Safety Programs Department as requiring both research and development.

Phase I, called the Ford Employe Skilled Driving Program, was designed as a research project to test four basic hypotheses, using company employes as participants.

- The performance of experienced drivers can be improved through training and practice.
- Driver performance can be measured.
- Measurable improvement can be accomplished quickly and economically.
- Improved driver performance will reduce accident probability.

Later phases will test these hypotheses in the field, using both fleet drivers and participants from the general driving population.

To gain a higher level of objectivity and the advantage of specialized expertise for the Phase I research project, the Traffic Safety Programs Department negotiated service contracts with the Highway Safety Research Institute (HSRI) at the University of Michigan and with the American Academy of Transportation (AAT) of Ann Arbor, Michigan.

Under its contract, HSRI agreed to perform the following:

I. Define the experimental plan.

II. Coordinate with and advise Ford project manager throughout the program.

III. Prepare a procedural manual covering items I and II above.

IV. Receive data from the Ford project manager and AAT, obtain driver data from the Michigan Secretary of State, construct a data file, and analyze the data, subjecting them to all practicable statistical techniques.

V. Write a final report, and make oral presentations as required.

AAT contracted to reduce paper tape recorded data from vehicle installed driver performance instrumentation and to provide the services of six persons who could be trained to serve as instructors under the direction of Phil Gram, Ford project manager.

This is the final report prepared by HSRI on this driver research project.

Fletcher N. Platt
Traffic Safety Programs Manager

ACKNOWLEDGMENTS

This evaluation was possible only through the assistance of a sizable group of people.

While at HSRI, Dr. William L. Carlson, now Associate Professor, St. Olaf College, Minnesota, was jointly responsible with the author for the experimental design and made the initial recommendations for analysis procedures. James O'Day, Head, Systems Analysis, HSRI, was Project Director, and most helpful throughout the evaluation. Dr. Richard Kaplan, Research Psychologist, HSRI, and Dr. Donald Smith, Michigan State University, assisted in test development. Dr. Jairus Flora, Assistant Professor of Biostatistics, consulted on statistical techniques, and carried out several of the analyses on the Perception of Hazards and Unusual Uses tests, as well as on the follow-up periods for crash data. Dr. Samuel Schultz II, Associate Professor of Psychology in Nursing was also a valued consultant on statistical techniques. Dr. Arthur C. Wolfe, Senior Research Associate, HSRI, advised on survey design. Mr. John Mahone of HSRI built the data files and carried out many of the computer analyses.

To Fletcher N. Platt, Manager, Traffic Safety Programs, Ford Motor Company; to Mr. Phil Gram, Manager of the Program (now Driver Safety Programs Manager) and his staff, and to the instructors, the author is especially indebted for their painstaking responses to our requests to execute many difficult procedures necessary to maintain the integrity of the experimental design.

Several other people at Ford Motor Company and the American Road Insurance Company were very generous with their time in facilitating our use of the various types of insurance data. These included Mr. Grant Wood, Mr. Anthony Matzdorf, and Mr. Robert Ingersoll.

CONTENTS

CONTENTS (Continued)

TABLES

TABLES (Continued)

FIGURES

1.0 INTRODUCTION

1.1 CURRENT KNOWLEDGE ABOUT DRIVER IMPROVEMENT AND RETRAINING

Programs to improve the ability of adults to drive automobiles have been attempted in many countries and in many parts of the United States during the last two decades. Most of these have sought to change the behavior of drivers who are relatively over involved in accidents or violations. To try to change those apparently most in need of help seems reasonable; however the assumption is implicit that a small percentage of (recalcitrant) drivers is responsible for the large majority of accidents.

There undoubtedly is a small percentage of drivers whose accident and violation experience is abnormally high and remains so over a period of some years. However, the contribution of this group to the overall accident problem is not great. Variations in record-keeping and enforcement practices result in differing estimates of their contribution, but studies have consistently shown that even *removal* from the population of drivers with accumulations of accidents or violations during a two- or three-year period would have little effect on accidents in the following two or three years. For example, Campbell (1972)[1] suggests that removal for two years of all North Carolina drivers with three or more violations in the previous two years, would "prevent" only 3.8% of the accidents in the later period. Similarly, Burg (1970)[2], reports that removal of all drivers with two or more accidents in a three-year period would eliminate 3.9% of the drivers and only 8% of the accidents. Rehabilitative programs are clearly warranted for such drivers, but there is much justification for efforts to upgrade average drivers.

Unfortunately, the belief is widespread that only the deviant drivers should be re-trained--either because they "deserve it" and/or because average drivers would find re-training

unpalatable or even insulting. Thus most of the data available from systematically evaluated driver retraining or improvement programs relate to rehabilitative settings. While these studies may yield indications of the usefulness of improvement procedures, the procedures normally involve visible or implied compulsion by the courts or driver licensing authorities, and it is difficult to generalize their findings to drivers with unexceptional driving histories. Moreover, some of these studies have reported improvements in drivers which very probably would have occurred whether or not they had been "treated".

Evaluations of other adult driver programs fall primarily into two groups. The first group comprises studies of programs for professional drivers, such as truckers and taxi cab drivers. Once again, these can scarcely be generalized, especially where drivers are highly selected, and where a culpable accident may result in company sanctions, perhaps dismissal. Even the experimental American University/U.S. Coastguard program (1972)[3], although aimed at private driving, may be operating under something analogous to fleet pressures.

The second group of studies cover programs offered to the general driving public, such as the Defensive Driving Course, but few of these employ study designs which are capable of yielding useful results. Furthermore, most such programs with open enrollment attract segments of the driving population which may be far from "average". Again, there is little doubt that this constitutes a target group worthy of appropriate driver improvement efforts; but we must look elsewhere for the information needed to efficiently upgrade large, unexceptional populations of drivers.

The ideal experimental program is easier to describe than to attain. It should certainly be carried out under an experimental design, rather than an observational or survey research design. It should avoid complexity in the content of training courses so that the groups to be compared under the design are

as large as possible (for statistical purposes), and so that some questions about amounts of training may be addressed. Such a program needs a large sample which is highly representative of the general driving public, and which is randomly assigned to several treatment or control groups, preferably after stratifying on key sociological variables. It should achieve a balance between the amount of persuasion needed to ensure high participation and that amount which appears to bring coercion to bear directly upon the participant's driving habits. Finally, the objective of the experiment should not be to prove or disprove that the program is "the answer" to the accident problem; rather it should explore the feasibility of a particular setting for a program, together with the nature of its successes and failures with unexceptional drivers.

A large corporation, whose employes depend almost exclusively on private cars to travel to work, recommends itself as a setting for a number of reasons. Among these are the possibility of management support to get high participation from the gamut of employes, good communication channels, and some data-gathering opportunities not practicable in "general public" settings. Unfortunately, companies interested in driver improvement are not usually prepared to adopt an experimental format for programs, and in particular are reluctant to give their support to the simultaneous adoption of several re-training procedures.

However, Ford Motor Company has for some time undertaken driver research, and in recent years it has applied this to the development of driver training aids and materials. Out of this, and from growing concern for off-the-job automobile accidents as well as losses to cars leased by Ford to certain employes, the company embarked in 1972 upon a long term effort to develop and disseminate adult driver re-training methodologies aimed, initially, at company employes. The first part of that program was an experimental "Employe Skilled Driving Program" to investigate some of the basic questions about

re-training large populations of unexceptional drivers. This report is a technical discussion of the evaluation carried out by HSRI of that program.

1.2 ORGANIZATION OF THE REPORT

The remaining five sections provide a systematic review of the Ford Employe Skilled Driving Program.

The operational aspects of this experimental program are discussed in a separate section (2.0) from that dealing with the evaluation methodology per se (3.0). A full section (4.0) is devoted to the nature of the data base, together with such validation of sampling and operational assumptions as are necessary to delimit the generalizability of the results. Detailed conclusions from the data are discussed together with the presentation of results in Section 5.0. A brief concluding section (6.0) gives several recommendations for future work in this area.

2.0 DESCRIPTION OF THE FORD EMPLOYE SKILLED DRIVING PROGRAM

This section gives a brief account of the training program, discusses the experimental format of the program in relation to the population from which subjects were drawn, and describes the operation of the program.

2.1 PROGRAM CONTENT, TRAINING METHODS*, INSTRUCTORS, AND FACILITIES

The content of the Ford Employe Skilled Driving Program was developed primarily from the more advanced training materials produced by the company in recent years. It was designed to encompass a number of aspects of driving ability which could be covered at different levels of detail in the instructional procedures.

The main theoretical bias implicit in the selection of content, is that adults are served better by efforts to improve their ability to think systematically about essential information while driving, than by attempts to drill them in specific procedures for traffic maneuvers. Central to the teaching included in this program is a paradigm for analyzing traffic situations known as the "Decision Pattern". This is closely tied to other selected principles which may be generalized to most normal driving conditions; these include headway timing, maintenance of a space cushion, and associated visual habits. Specific driving maneuvers and highway traffic situations are included to exemplify the application of these principles. In addition, a moderate amount of emphasis is given to emergency and abnormal driving conditions; and general safety habits, such as belt usage,

---

*In order to maintain an independent evaluation, HSRI was not involved in the development of the instructional program. It was assembled by Ford Motor Company with the assistance of its driver education consultants. An outline of the syllabus for the training sessions, and a list of available individual study materials will be found in Appendix 1.

are encouraged.

Teaching methods fall into two categories: formal training and self-instruction. (The distribution of these among treatment groups is discussed later.)

The two-hour formal training sessions were equally divided between classroom and behind-the-wheel components. On each occasion, the classroom component combined lecture-discussion with film and filmstrip presentations. Two of these sessions also employed a simple programmed learning device which presented a multiple choice test with the aid of a filmstrip. The behind-the-wheel components provided a car and instructor for every two participants. On each occasion, every participant drove for 20-25 minutes under instruction relating to the classroom teaching of that day. Standard routes, in the vicinity of Dearborn, Michigan, were selected for their mix of driving situations, and were varied from session to session.

Self-instruction was used in two forms. Firstly, a course of home instruction covered the content of the program with booklets and quizzes. Secondly, a library was set up adjacent to the classroom. This was equipped with consoles for individual use of tape and filmstrip modules covering the topics in the training program and other instructional materials developed by the company. Take-home booklets corresponding to the instructional modules were also available.

The instructors for this experimental program were purposely selected from varying backgrounds. The rationale for this was that a program designed for large populations would inevitably operate without highly trained instructors.

Six instructors were initially employed. They included two women, aged 46 and 28, and four men, aged 50, 28, 24 and 23. The 24-year old man was black; all other instructors were white. During the course of the program, the 24-year and 28-year old men left, as did the 28-year old woman. Two additional instructors, men aged 27 and 23 were engaged as replacements. Their

backgrounds included carpentry, psychology, civil service clerical, engineering and filmmaking. None was trained as a teacher, although several had some kind of teaching experience. One was especially experienced in counselling. Driving instructors of any kind, or police drivers, racing drivers and other specialized automobile drivers were specifically excluded in the attempt to create a realistic trial of this kind of program.

Special classroom facilities were constructed for the program in the lobby of Ford's World Headquarters Building in Dearborn, Michigan. All training and testing sessions were based at these facilities.

2.2 EXPERIMENTAL FORMAT

The primary purpose for experimentally evaluating the program was to yield information on the feasibility of re-training large numbers of adult drivers, and in particular, to examine the relative effectiveness of several levels of intensity of training within a company setting. This required a substantial sample of employes, preferably drawn from a wide range of age and socio-economic levels. It was decided to draw subjects from the gamut of salaried employes. Ford Motor Company has approximately 10,000 salaried employes in the Detroit metropolitan area; about 2,500 of these are managerial, while the remainder are on the general salary roll.

In order to distribute the sample more evenly over the entire salary range, approximately equal numbers of managerial and general salaried employes were randomly selected by Ford to be sent invitations to participate in an innovative advanced driving program. Figure 2.1 shows that out of the 10,000, 954 general salaried and 1,046 managerial employes were sent a letter of invitation; 400 of the general salaried and 602 of the managerial employes accepted. The acceptance was high enough to divide the volunteers into four treatment groups, giving three levels of training intensity and a control. The highest level (8-Hour group) received four two-hour training sessions; next in

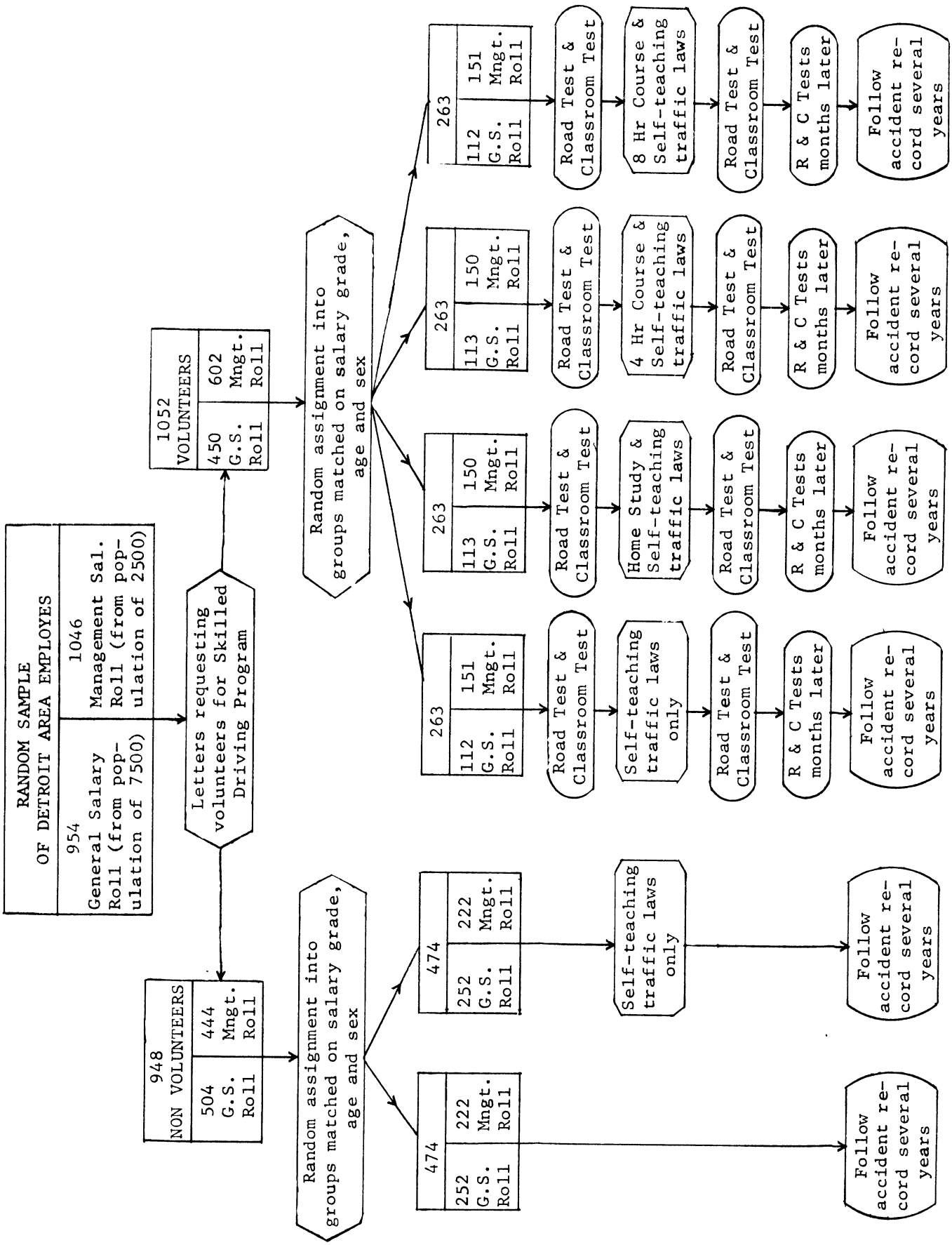


FIGURE 2.1 The Ford Employe Skilled Driving Program Experimental Format

intensity was the 4-Hour group, who received two such sessions; the lowest level was the Self-Teach group, who were given a package of home study materials at the end of the pre-test sessions. All four groups, including the control group, were treated exactly alike insofar as they underwent the same sets of pre-tests and post-tests. Only at the end of the pre-test session were subjects told which of the four groups they had been randomly assigned to. For any subject, all training activities took place in working hours during a five-week period after pre-testing, at the end of which the post-tests were administered.

Because of the amount of contact the control group thus had with the program, and incidentally with other participants in the company, it was not reasonable to regard them as a classical "hands off" control group. Therefore, this group was given a new edition of the Michigan state driver handbook with a written exhortation to study it, and to complete a short quiz which was attached. (This handbook is entitled "What Every Driver Must Know"; the control is referred to hereafter as the WEDMK group.) With this type of control group, differences measured between the groups could be considered to be controlled for any Hawthorne effect which may occur in a company setting.

In addition to being oriented to their respective instructional treatments, all groups were told at the end of pre-testing that program participants had exclusive access to the audio-visual library*. The rationale for the library was that by making all of the instructional content accessible to all groups, measured differences in learning between the groups could be attributed to the training activities, rather than to variations in their access to the informational content. A record was kept of which library materials were used by each participant who took advantage of the facility.

---

*An additional facility was set up at a local public library later in the program for general use. Records of library usage revealed that an insignificant number of program participants used the remote facility.

The sampling procedure used to obtain the 2,000 invitees from 10,000 salaried employes was simple random selection from computerized salary roll records, with a weighting factor to increase the selection of managerial employes.

Assignment of volunteer groups was completed using strafified random sampling. Previous driver research has emphasized the importance of controlling for major biographical factors, such as age, sex and socio-economic status. In this study, salary grade, in the form of a scale from 1-18, was considered to be an adequate surrogate for S.E.S. The 1,052 volunteers were stratified on salary grade, age and sex, and then randomly assigned to the four treatment groups.

A third set of tests, the "follow-up", were administered about four months after the post-test. To select follow-up subjects, two simple random samples of about 25% of all pre- and post-tested participants were defined. These samples were listed separately in ascending order of two stratifying variables (age within a salary grade). Subjects on the first list were telephoned; if they agreed to take part in the additional tests, the next subject on the first list was telephoned; if they refused, the subject with the same number on the second list was sought as a substitute. Whether that person accepted or not, the next subject to be telephoned was from the first list. This procedure yielded as many follow-up subjects as could be accomodated in the project, about 15% of all participants, or approximately 20% of those pre- and post-tested. It was originally intended to follow-up test 50% of those completing the program; sampling assumptions for these smaller samples were tested and are discussed in Section 4.2.

In addition to the administration of tests, the experimental format also provides for the collection of additional biographical and opinion data at pre-test, and for a mailed survey requesting feedback from program participants four months after their post-test took place (whether or not they attended it).

Two kinds of driver record data were also collected for all volunteers. Firstly, the official Secretary of State (S.O.S.) record of accidents and violations was obtained for a two-year period ending with the month prior to completion of the program for 710 of the 750 volunteer participants. Of the remaining 40, 28 had no Michigan driving record (mainly accounted for by those resident in Ohio or Ontario), and twelve declined permission. Secondly, for managerial employes, Ford provides leased cars as a fringe benefit. Crash data is available for these cars, as they are insured for physical damage by a Ford subsidiary, The American Road Insurance Company (TARIC). These data were collected for approximately 400 managerial participants over the same periods as the S.O.S. data. Both types of driver record data will be periodically examined in the future to investigate the effect of the program on accident and violation rates.

For the 948 invitees who did not volunteer, it was unfortunately not possible in this phase of the long term effort to provide any training activities. However, as shown in Figure 2.1, they were randomly assigned into two groups, after stratifying on the same basis as the volunteers. One of these groups was mailed the "What Every Driver Must Know" booklet and the accompanying letter and quiz, as given to the WEDMK group; these were mailed at intervals to randomly selected sub-samples over the period June to November 1972. This was to permit the eventual comparison of driver records between this group, the remaining non-volunteers who received nothing more than the original invitation, and the volunteer groups. Although of less importance than the volunteer group, in the long term this does provide an additional treatment analogous to no-threat warning letters in driver improvement studies. For these purposes a simple random sub-sample of 124 of the 948 non-volunteers was selected using the final digit of the subjects' Social Security Numbers; the previous S.O.S. driving record data was collected for the 117 out of the 124 who had a valid record in the State of Michigan. For comparability, it was necessary to distribute the final months

of the two-year periods for which driving history was examined throughout the period during which volunteers completed their training; differences in driving records might otherwise be biased by historical artifacts, such as excessively severe winters. Final months were randomized over the period June-November 1972 so that non-volunteers who were mailed the WEDMK package could be considered to have completed training shortly after receiving it. By comparison, the final months of the volunteer previous driving record periods range from May 1972 to March 1973.

An additional group to be considered are the 302 volunteers who did not take part in the program when it was offered. By a procedure identical to that used for non-volunteers, 103 valid Michigan driving records were obtained out of 109 selected from this group. In this instance, the final months of the two-year periods were randomly distributed over the same period as that covered by the volunteer records.

2.3 PROGRAM OPERATION

The program was conducted between mid-June 1972 to the end of April 1973. Originally, it was intended to finish training by mid-December 1972, but partly because of some unusual demands on salaried employes during the late Summer of 1972, many participants had requested that they be re-scheduled later in the year. Many requests could not be accommodated, and therefore it was decided to break from Thanksgiving until the new year. The program could not be completed until the end of April, partly because of the time lag between pre- and post-tests, which was held constant at five weeks.

The participants were scheduled into a total of eight overlapping "cycles". Each cycle consisted of an eight-week period within which up to forty participants from each of the four treatment groups could be processed as far as the post-test. Within each cycle, participants were assigned to four sections, which approximately corresponded to salary grade quartiles, for

testing and (where applicable) training sessions. Although for logistical reasons each treatment group was always trained and tested on the same two days of the week, section order was randomly re-assigned from cycle to cycle. Pre- and post-tests were always arranged for a given section on the same day of the week and at the same time of day, to try to keep traffic conditions as constant as possible. A great deal of re-scheduling, both within and between cycles prevented the maintenance of complete randomization in all these factors.

Instructors were randomly assigned in both their teaching and their testing roles. These assignments were sometimes changed because of normal operational problems.

3.0 EVALUATION DESIGN

This section comprises a brief discussion of the model used in this evaluation, and a description of the instruments which were selected.

3.1 EVALUATION MODEL*

The conventional response to the evaluation of programs intended to change driver behavior is to try to measure the ultimate criterion for success--accident (and sometimes violation) experience--very directly. Thus, driver records before and after the program are compared. However, for a number of reasons, observed reductions in accidents tend to be either insignificant or spurious.

Statistical significance at a given level is a function of the size of the mean reduction and of the number of people in the sample used. A great many factors contribute to accident rates, and therefore a course of instruction alone cannot reasonably bring about a large reduction. Typically, statistically significant reductions require much larger samples or longer periods than are convenient for carrying out evaluations. Even if accident data are collected over several years, the population from which the sample is drawn may undergo some important changes, not the least of which is normal job and residence turnover for people of working age. Moreover, all uses of accident data are open to very serious biases in accident reporting and record-keeping.

Statistically significant accident reductions have sometimes been attributed to adult driver programs. However, most adult programs have been directed towards people who have recently experienced a poorer-than-average driving record, and regression to the mean ensures that such groups will improve in any case.

---

*A detailed argument for the evaluation model applied to this program will be found in Lee (1973)[4].

If accident reduction is subject to many influences besides the introduction of a driver retraining program then it is reasonable to try to measure the impact of a program more directly. Indeed, it is reasonable to consider a spectrum of behavioral effects which range from those directly attributable to the program, such as the fact of attending a course, to the scarcely detectable influence of the same course on something as complex as an ultimate reduction in accidents. In fact, immediate effects such as attendance figures have been used in many safety programs as a criterion of success. However, this immediate "monitoring" of the program, even if it accompanies the use of accident records, is far from a complete evaluation. For this reason, "intermediate" effects have gained considerable attention in recent years. Inherent in the notion of measuring intermediate effects is that the connection between the application of a driver program and an ultimate change in accidents and associated losses is via some changes in individual drivers (Figure 3.1)*.

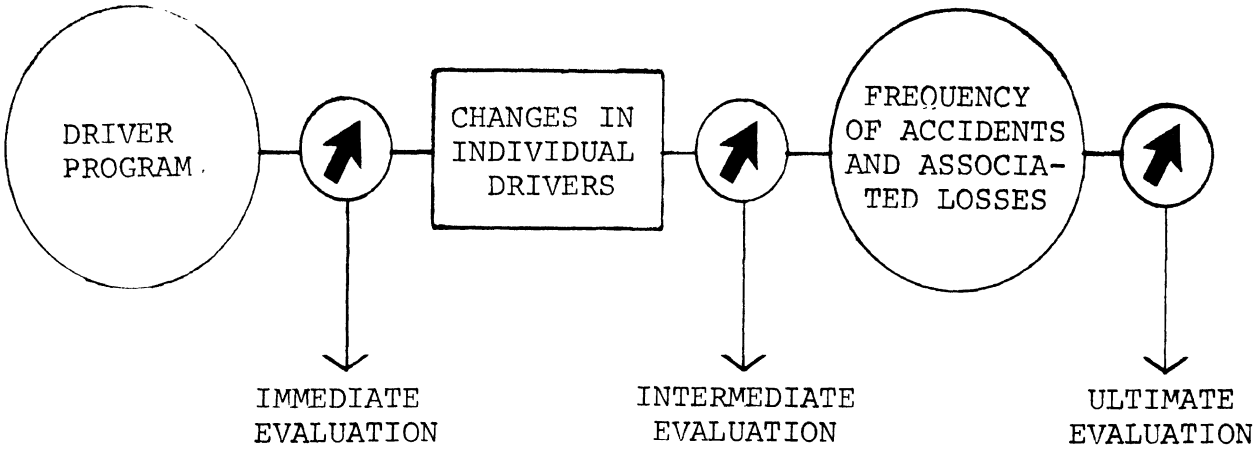


FIGURE 3.1 Basic Classification of Types of Evaluation For Driver Programs

*Intermediate measurement, as placed conceptually between the immediate monitoring of program activities and ultimate measurement using accident data, should be distinguished from the time frame of "short- medium- and long-term measurement".

Examples of techniques for measuring these changes are on-street driving tests and driver knowledge examinations. Although various forms of these have, of course, been in use for many years, the overall state-of-the-art of intermediate measurement is very poor, despite substantial effort to improve it. After the well known Federal Highway Safety legislation of 1966 explicitly endorsed the trend of substantial public investment in various kinds of driver instruction, research funds from the same legislation were allocated for four fairly costly explorations of the problem of evaluating driver education and training (1968)[5]. A fifth contract (1969)[6] provided a plan for a substantial program of further research, starting from a very detailed task analysis of driving, and involving the development of numerous tests and measures. The plan has greatly influenced the awarding of Department of Transportation research contracts in this field since. The task analysis was completed in 1970, and by the end of 1972, substantial progress had been made in driver knowledge testing; but in most aspects of performance measurement, progress has been disappointing. No single test exists which, even with much experimental control, can provide a definitive evaluation of a driver training program in a period as short as six months or a year. Yet a decision whether to retain or expand a program is usually needed long before adequate accident data can be gathered. Therefore, some paradigm is needed to select from the many variables which are measurable, and which might be suitable as surrogates for the ultimate criterion of success--accident and loss reduction.

The concept shown in Figure 3.1 can be operationalized and expanded as a "causal chain" between the application of the program and its ultimate influence on the accidents. The causal chain model was explicitly applied to a wide range of highway safety program evaluation problems in an earlier HSRI study (O'Day, et al., 1971)[7]. The model was developed further in the present study to define a series of stages at which changes in drivers might be measured. (Figure 3.2).

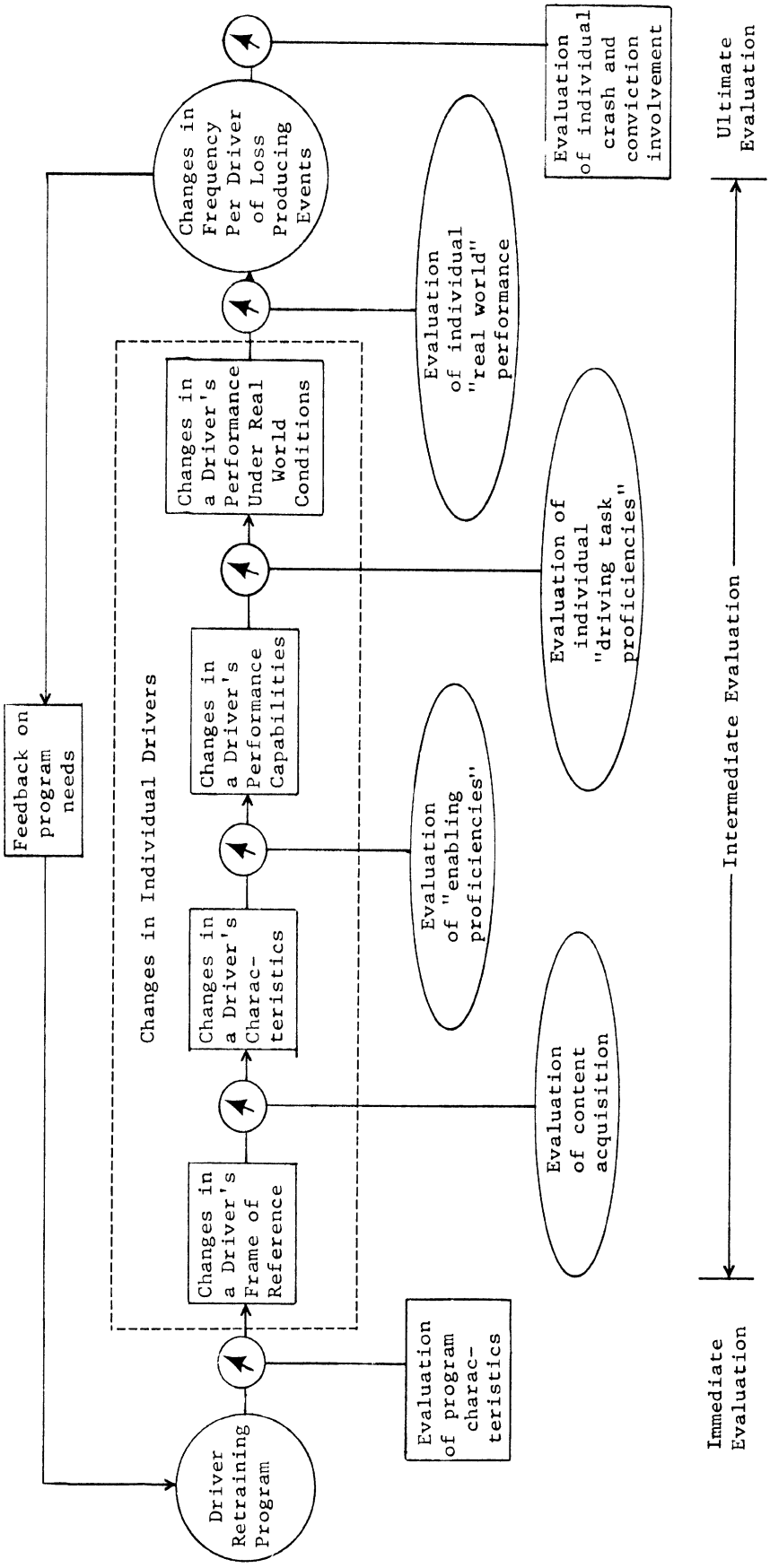


FIGURE 3.2 A Development of the Causal Chain Model for the Selection of Intermediate Instruments

The first stage, changes in a driver's frame of reference, merely considers the obvious requirement that the informational content of the program was understood. For this, content acquisition tests, usually in the form of written quizzes on program content, are appropriate.

The second and third stages are both defined in terms of what a driver is capable of, rather than what he necessarily does. His "characteristics", as used by the second stage of the model, refer to a set of basic proficiencies which enable him to operate as a driver. It is at this point that the most prevalent model for driver testing is considered, namely the cognitive-affective-psychomotor trichotomy of behavior (usually interpreted as knowledge-attitudes-skills). Separate tests of these are valuable, but not sufficient for a comprehensive evaluation. Relevant techniques include tests of general driving knowledge, simple and complex reaction tests, and vision tests. Despite the undoubted importance of a wide range of social and motivational factors to driving behavior, attitude tests are generally disappointing, particularly with unexceptional drivers.

The third stage considers the driver's capability to perform the major subtasks which comprise driving. Most classifications of these tasks are of the input-process-output variety, and laboratory settings for testing each of the three can be achieved. Examples include: for input, the ability to perceive hazards in films or slides of traffic situations; for process, the ability to make rapid decisions about verbally or visually presented critical situations (as used in some of the televised "driving tests"); for output, the ability to control a vehicle under demanding conditions at an off-street test facility, or to communicate intentions in hazardously ambiguous situations.

Finally, at the fourth stage, changes are considered in the extent to which performance is affected under real world traffic conditions. There are events occurring at the rate of several per minute or more which it may be useful to measure. The input-process-output model may again assist in selecting specific

behaviors to be observed (for example, visual scanning, appropriateness of decisions, and dynamic vehicle positioning), but there are also questions of timing and sequencing under real traffic conditions. Tests which measure traffic conflicts rather than the driver's individual actions, comprise a promising form of measurement which is meaningful at this stage of the causal chain, although most work in this area has concerned highway locations rather than drivers. The most practicable form of test for this stage is the on-street observation of driver actions by an observer or an automatic recording device present in the car. Less obtrusive forms of observation (from a trailing vehicle for example), are very desirable but rarely feasible.

At all four stages in the causal chain model, much confusion can arise over the nature of the criteria for success which are erected with each type of test. An important distinction can be made about the way in which particular outcomes are defined as successful. On one hand, outcomes may predict, or have high logical relevance to the ultimate criterion of success, namely accident reduction; on the other, success can very credibly be defined as the attainment of instructional objectives, which themselves may only be assumed to relate to accident reduction. At the present state-of-the-art of testing, the latter definition of success must assume major importance, even under the "real world" test conditions.

Given this broad range of measurement possibilities, a decision must be made on the number and mix of instruments to be employed. There is, in general, a cost choice to be made between increasing the amount of experimental control to improve the fidelity of test results, and increasing the amount of testing itself. However, even at the expense of some confidence in individual test results, *it is important to select instruments from as many stages in the causal chain as possible.*

This ensures that a maximum of information is yielded by the evaluation. The use of only one or two types of measurement can

be close to catastrophic if unforeseen problems arise with the experimental controls. Under ideal circumstances, each of the four stages of intermediate measurement would be fully utilized together with immediate and ultimate measures, providing the opportunity to trace in great detail the effect of the training throughout the causal chain. While this may not be fully attainable at present, the model provides a highly preferable alternative to evaluations built around some independent a priori definition of "safe driving".

3.2 DESCRIPTION OF INSTRUMENTS

Immediate Measurement

For the purpose of monitoring the operation of the program, records were routinely kept of attendance and instructor assignments at all test and training sessions, of library usage, and of reasons, where available and appropriate, for non-completion of the course.

In addition, two attempts were made to measure participants' opinions of the program. The first consisted simply of a yes/no question about whether they thought Ford should offer skilled driving instruction for its employes, together with two open-ended questions on the reasons for their answer, and for taking advantage of the course. These were embedded in a questionnaire administered at pre-testing. This was primarily intended to gather biographical data, not available from company files, but which was needed to validate sampling assumptions (such as driving experience, and previous driver improvement course participation). A copy of the questionnaire is included as Appendix 2.

The second attempt to measure opinions was through a survey which was mailed to all participants approximately four months after their post-test date (whether or not they completed the post-test). This included Lickert type items on the usefulness of the program, their approval of skilled driving as a company offering, and ratings of their own driving ability before and after the program. Other multiple choice questions asked whether

and when they would like to re-take such a program, whether they would like it offered to members of their families, and who they believed to be responsible for the largest number of all automobile accidents. Open-ended questions invited comment on the most and least useful parts of the program, instances from their driving experience of how it helped or confused them, why they took part, and upon changes or additions they would recommended for any part of the program. Appendix 3 comprises this survey.

Finally, at post-test, participants were invited to write down the number of near accidents they had recorded on a special score card which was handed out at pre-test. This is an educational device, rather than a criterion instrument, but the reporting rate of each group is useful as an index of participant interest and can thus be regarded as part of immediate evaluation.

Intermediate Measurement

At least one test was used at each of the four stages discussed in Section 3.1. A content acquisition test consisted of thirteen multiple choice questions on the information content of the program. It was presented as Parts II and III of a written test which, to minimize contamination from discussion and recall, was not administered at pre-test. Parts I and IV of the same test contained another seventeen multiple choice items, which provided a test of knowledge about driving techniques and problems (per "enabling proficiencies" in the causal chain model). The written test as a whole was developed by Ford from a larger set of items which were pre-tested for difficulty on Ford employes not involved in the program. A copy is included as Appendix 4.

Two instruments were chosen to test "driving task proficiencies". In the area of "input" subtasks, an attempt was made to build two comparable, short tests to measure the Perception of Hazards by drivers. The items used were suggested by an instruction booklet from Illinois State University (McPherson and Cooper, 1966)[8] which describes how to construct such a test from

a filmstrip of the same name published by the Shell Oil Company. Each frame of the filmstrip shows a potentially hazardous traffic situation. These were shown for five seconds, after which subjects were read a list of three to five clues to danger, at least one or two of which were present in the preceding frame. Subjects answered true or false to each clue, and were scored on the sum of their correct responses over six frames. The instructions for the test suggested weighted demerit points for wrong answers. The suggested weights, and also simple demerit scoring (one negative point for each incorrect answer) were tried on the accumulated data, but no advantage was found to these more complex scoring methods. Both versions of the test were automated using a filmstrip and tape cassette. It was not possible to fully pre-test the versions for comparability owing to the severe time constraints at the start of the program.

Two versions were needed because it was considered desirable to make pre-test-post-test comparisons, but to administer the same version twice would again invite contamination for those participants who discussed its contents between pre- and post-test. The versions were randomly assigned to approximately equal numbers of pre-test sessions; at post-test, participants took the version they were not assigned at pre-test.

For "process" subtasks, a simple creativity-testing technique was applied, somewhat speculatively, to the thinking of drivers. The technique, known as "Unusual Uses", is to ask a subject to name as many uses as possible for a given object; in this case, participants were asked to write down as many actions as possible to maintain safety in a hazardous driving situation, which was described*. Two such situations were assigned on the same random split-half basis as was used for the Perception of Hazards test. Scoring was by summation of the number of mutually exclusive responses. A copy of the test comprises Appendix 5.

*This technique was suggested, and the items written, by Dr. Richard J. Kaplan of HSRI.

Testing of performance under real world traffic conditions was restricted to individual driver actions; conflict measurement was not attempted. A standard fifteen mile route was specially designed to include a warmup period on service roads, three freeway stretches, each about three miles in length, and three urban driving situations, each including a congested left turn from an arterial into a residential street. One freeway section was followed by all three urban/left turn sections, followed by the remaining two freeway sections. During the second urban/left turn and the second freeway sections, participants were asked to perform a secondary task to simulate the more stressful conditions which occur during distraction. The tasks used consisted of reciting the alphabet forwards or backwards at given time intervals using the dashboard clock. This provided one section each of freeway and urban/left turn driving under normal driving, one each under stress, and one each under recovery from stress. For both types of driving, the stress and recovery sections were contiguous.

This test route was designed to accommodate two methods of measurement. Firstly, an on-board automatic recording device, designed by Ford Motor Company, provided an Objective Behind the Wheel (OBTW) Measure. The output from this device was a continuous graphical record on paper tape of steering reversal rates, speed, and brake applications over time. In addition, over each of the freeway sections, a formula of speed and steering reversals was computed and displayed as an index number. This was manually cued at set highway locations. The number was scrambled in a manner unknown to the observers to avoid contaminating the subjective scoring.

The device is a modification of instrumentation which has been developed since the late 1950's in Ford Motor Company's program of driver and highway research. An overview of its development is given in Platt, Gram and Hobday (1969)[9], and an earlier but more detailed discussion of its application will be found in Platt (1964)[10].

The paper tapes for each test subject were transcribed* into digital form. Variables recorded comprised steering reversal rates over each of the six sections of the route, total time elapsed on route, the index numbers and average speed for each of three freeway sections, and total brake applications. In addition to these variables, various combinations were used in analysis, including ratios of stress to pre-stress data. More detailed material on the objective instrumentation will be found in Appendix 6.

Secondly, a Subjective Behind the Wheel (SBTW) test was designed to measure group mean differences in the application of principles taught in the program**. The route was used to provide three replications (corresponding to pre-stress, stress and recovery) of three observations of freeway driving, and of six observations during the urban/left turn sections. Each of the nine items were developed by HSRI in cooperation with the instructors, who were to perform the observations. It was especially important to achieve and maintain consensus on the interpretation of the items, as each item required the observers to score on a yes/no basis whether a number of related behaviors, taken together, were performed satisfactorily. The items all relate to the instructional content of the program, and are logically related to accident-free driving. A score sheet with key-phrases, and a full-length version of the items, will be found in Appendix 7. In addition, during the freeway sections, an estimate in seconds was made of the amount of headway consistently accepted.

Subjective observations were not made continuously. Rather, a one to two minute sequence was defined between fixed

---

*The transcription was completed under contract with Ford by the American Academy of Transportation, Ann Arbor, Michigan.

**This test was developed by Martin E. Lee with assistance of Dr. Donald Smith, Driver Behavior Project, Highway Traffic Safety Center/Department of Psychology, Michigan State University, East Lansing.

landmarks within each of the six sections of the route. These were carefully chosen to include maximum comparability between each of the freeway sections and between each of the urban/left turn sections. A short but intensive training period was conducted by HSRI of the initial set of instructors in the operation of this test, following immediately upon the finalization of its content. All possible pairs of instructors jointly rated Ford employes who volunteered as guinea pigs, (but who were not part of the program). Differences in scoring were discussed after each test run and inter-observer reliability rose considerably as a result. Thereafter, this exercise was periodically repeated to try to maintain maximum consistency between the instructors. On two occasions, November 1972, and March 1973, item analyses were carried out from such a series of runs, and the results are included in Appendix 7. Inter-observer reliability was understandably not as high as would be attained with simpler tests; therefore, whenever feasible, the same observer was assigned to a subject for pre-, post- and follow-up tests. Also, instructors were not permitted to act as the test observer for a participant whom they had taught in any of the road training sessions. All analyses comparing two administrations of this test on the same subject were confined to cases for which the observer remained constant. It was intended that at all times during the program, the treatment group membership of test subjects would be concealed from the observers. Partly because of scheduling regularities, which were essential for logistical reasons, it was impossible to maintain this condition.

In the time available, it was not possible to establish test-retest reliability experimentally. However, within the experimental design, the replication of the test for the control (WEDMK) group gives some indication of this, subject to some possible drawbacks, such as differences in Hawthorne effects. For the 428 subjects who were tested by the same observer, pre-test score correlates highly significantly with post-

test score at r = .5435.

Because of a tendency for evaluations of driver training programs to reflect major biographical characteristics rather than training effects, a study was undertaken on the first five cycles of data to establish the relationship between this test and age, salary grade, and driving experience. Regression models including all three factors were developed on three samples of the subjects; none of the models could explain more than 7.3% of the variance in pre-test sum score. An edited version of this study is included in Appendix 7.

Ultimate Measurement

Two types of accident or loss data are available for the long-term comparison of treatment groups. Both have already been used to compare two years of driving history prior to entering the program.

Under a standing arrangement with the Secretary of State, official Michigan driving records may be accessed by HSRI for strictly confidential research purposes. Group accident and violation rates may, however, be published. The number of all collisions, convictions and points, are being stored, as noted elsewhere, for 95% of the volunteers, 35% of the cancellations, and 12% of the non-volunteers.

For those volunteers eligible for leased cars as a job benefit (approximately 400), detailed data are available to this evaluation from The American Road Insurance Company (TARIC), who underwrite damage to these vehicles. TARIC experiences around 0.7 claims per vehicle per year, but using a hand-sort method of claims records, matches were found at about 0.3 claims per vehicle per year for program participants. This is largely a function of the number of claims which are processed without an employe name, especially when minor damage is rectified at the time a car is turned in for replacement, or of the change in leasing number which results when an employe changes

his company location; TARIC data is only stored for this evaluation if the name and lease number on the claim sheet matches that of the participant. Information recorded consists of the total number of collision claims, total non-collision claims, collisions involving a liability claim against the employe, claims involving subrogation attempts on behalf of the employe, and total dollar value of damage to the leased car. Unlike the SOS data, a record only exists for those with accidents on record. Therefore, claim and dollar rates have to be computed on the basis of all participants who are eligible for leased cars. A problem with these otherwise very promising data is that the driver of the leased car at the time of collision is often unknown; moreover, because the lease fee includes no-deductable insurance, and no other financial penalties for accident involvement, these cars receive disproportionate use by higher insurance risks, such as the teenage children of employes.

Two other types of insurance data were originally intended for use in this evaluation. The first was the dollar value of liability settlements against the drivers of leased cars. This proved extremely difficult to access in a consistent manner for individual drivers, because of a change in the Underwriter during the period for which data was sought*. Secondly, an arrangement was made, courtesy of Blue Cross in Detroit, to record instances of medical claims made by participants in respect of injuries sustained in automobile accidents. Unfortunately, these data are much more sparse than expected, and are not being placed in the data base.

---

*Thus the files contain no data for variables 302 and 308 in the codebook (Volume 3 of this report).

4.0 DISCUSSION OF DATA BASE

This section describes the data base which has been developed from the Employe Skilled Driving Program, and compares certain characteristics of the treatment group sub-samples, before and after attention caused by dropouts and bad data. Sampling and operational assumptions are examined statistically.

4.1 DESCRIPTION OF DATA BASE

The data base has been developed in the form of three computer files. The principal ("Participant") file contains data on the 750 individuals who participated, at a minimum, in a pre-test session. These data include biographical information from company files and from the questionnaire which was completed at pre-test, all test data (in raw form where this is appropriate), and survey responses from 402 participants*. It also includes two years of previous driving history data from Michigan Secretary of State records for 710 participants, and a similar period of insurance (TARIC) claims records for private vehicles leased from Ford by 187 of the participants. A further 207 participants were estimated to be eligible for leased cars, but for these, no claims record was found. Altogether, the Participant file has 432 variables; these are listed in Volume 3 of this report--"Codebook for Volunteer Participant File".

The two other computer files contain data, respectively, for the 947 Non-Volunteers, and 302 "Cancellations", the latter being those who volunteered, but for some reason did not then participate in the program. In addition, Secretary of State driver record data are included for the random samples of 117 Non-Volunteers and 103 Cancellations. The variables in both files correspond to

---

*Because the ESDP program concluded in April 1973, rather than December 1972, it was not possible to include survey responses received later than mid-July 1973. The file will be updated when all responses have been received.

variables 1-22 (biographical) and 312-322 (S.O.S.) in the Participant Codebook.

All three files have been built on IBM 360 hardware in the form compatible to the Highway Safety Research Institute's Statistical Research (SR) programs and the OSIRIS package developed by the Institute for Social Research. In addition, the Michigan Interactive Data Analysis System (MIDAS) may be used through a special interface.

4.2 VALIDATION OF SAMPLING ASSUMPTIONS

The experimental design calls for the comparison of four treatment groups of volunteers, and as an adjunct study, the comparison of a treatment and a control group of non-volunteers. It is important to validate the procedure used to assign employes to groups. A description of the procedure--stratified random sampling--is given in Section 3. It was expected that stratifying on age and salary grade would tend to ensure that the groups were comparable on other important variables. Unfortunately, there is always a certain amount of dropout from groups in programs of this kind, and it becomes essential to question the comparability of the groups after various levels of attrition. These levels are summarized in Table 4.1.

Using a one-way analysis of variance (ANOVA), a comparison was made of the mean values, by treatment group, of all available biographical and driving history variables, at the three main levels: those who were at least pre-tested (750); those who were both pre- and post-tested (560); and those whose subjective tests were conducted by the same observer at pre- and post-tests (428)*. Table 4.2 summarizes the means, and obtained ANOVA significance levels, for 24 variables. It will be seen that only two variables approach statistically significant differences.

---

*This last level deserves separate validation because of the necessity to use cases with matched observers exclusively in the analysis of the subjective behind-the-wheel test.

TABLE 4.1 Flow of Volunteers Through Program

TREATMENT GROUP:	8-Hour	4-Hour	Self Teach	WEDMK	TOTAL
Assigned	263	263	263	263	1,052
Pre-Tested	180	174	193	183	750
Session 1 Att'd	138	157	-	-	295
Session 2 Att'd	119	146	-	-	265
Session 3 Att'd	105	-	-	-	105
Session 4 Att'd	107	-	-	-	107
Pre- & Post-tested	120	162	144	134	560
Pre- & Post-tested By Same Observer	98	123	106	101	428
Dropouts	60	32	49	49	190
Cancellations	73	69	70	80	302

Both of these are S.O.S. accident variables in which recorded mean frequencies are extremely low--single vehicle collisions and injury collisions.

A further sub-sample which requires validating across treatment groups consists of those participants who were follow-up tested. Table 4.3 gives means values of selected biographical and driving history variables.

The follow-up tested sample contains, not surprisingly, some significant differences between treatment groups which have diminished to 28 or 29 in size. In particular, there are significant differences in mean age and in the variance of salary grade. Mean salary grade is also very erratic. The marginally significant differences in high school driver education experience can be attributed to age. Commuting mileage differences may not be meaningful with these small groups, but at least one outlier (130 miles) is contained in the WEDMK group.

As a whole, however, the follow-up tested sample with

TABLE 4.2

Comparison of Biographical and Driving History Data by Treatment Group at Three Levels of Attrition From Original Group Assignments

N's = Pre-tested: 750; Pre- & Post-tested: 560; Same Observer: 428

Variable Number	Variable Name	Sample	MEANS 8-Hour	4-Hour	Self Teach	WEDMK	ANOVA α
3	Cycle Number	Pre-Tested	3.9444	4.0052	3.7098	4.1039	.3330
		Pre & Post	3.9000	3.9444	3.5556	3.9179	.3667
		Same Obsvr	4.0510	4.2033	3.5660	4.0693	.1784
10	Salary Grade	Pre-Tested	10.650	9.4433	8.6477	9.2240	.1005
		Pre & Post	10.567	9.4877	8.6319	9.1418	.3146
		Same Obsvr	10.684	9.3984	8.8679	8.7228	.4016
14	Marital 1=married 2=single	Pre-Tested	1.1222	1.1701	1.1554	1.1202	.4214
		Pre & Post	1.1250	1.1975	1.1736	1.1194	.1987
		Same Obsvr	1.1122	1.1951	1.1698	1.1287	.3094
15	Race/Sex	Pre-Tested	not appropr.	not appropr.	not appropr.	not appropr.	.6369*
		Pre & Post					.6360*
		Same Obsvr					.5795*
18	Birth Year	Pre-Tested	29.067	30.072	29.430	30.437	.5116
		Pre & Post	28.267	30.086	29.549	30.194	.3553
		Same Obsvr	28.643	30.447	30.547	29.970	.4543
21	Year Joined Ford	Pre-Tested	57.967	58.428	57.461	58.339	.6739
		Pre & Post	57.350	58.654	57.576	58.216	.5454
		Same Obsvr	57.337	59.033	58.142	57.950	.5233
29	Freq. Drives Company Owned Car (5 point scale)	Pre-Tested	3.1073	3.0052	2.7884	2.8187	.1083
		Pre & Post	3.0336	2.9250	2.7660	2.7970	.4264
		Same Obsvr	3.0306	2.9587	2.8095	2.8900	.7255
32	Commuting Mileage Round Trip	Pre-Tested	26.575	27.820	27.440	29.186	.6658
		Pre & Post	24.412	26.938	26.896	28.410	.4229
		Same Obsvr	25.196	27.203	25.887	27.861	.7418
33	Annual Mileage (Thousands)	Pre-Tested	17.292	16.948	17.793	17.896	.5530
		Pre & Post	16.720	16.373	17.604	17.709	.2812
		Same Obsvr	17.289	16.492	17.264	17.455	.7133
34	Years Driven	Pre-Tested	25.268	24.608	25.085	24.355	.7954
		Pre & Post	26.398	24.500	24.855	24.478	.3629
		Same Obsvr	25.943	24.016	24.000	24.653	.4535
40	H.S. Driver Education (0=no; 1=yes)	Pre-Tested	.16867	.21649	.22404	.18579	.5252
		Pre & Post	.19091	.20988	.23704	.17910	.6707
		Same Obsvr	.19101	.21951	.28000	.18812	.3744
41	Other Driver Education (0=no; 1=yes)	Pre-Tested	.066265	.051546	.055556	.038251	.7033
		Pre & Post	.045872	.055556	.067164	.037313	.7220
		Same Obsvr	.056180	.065041	.060606	.029703	.6692
42	Prev. Driver Improvement (0=no; 1=yes)	Pre-Tested	.029940	.036082	.032787	.021858	.8728
		Pre & Post	.027273	.043210	.029630	.014925	.5596
		Same Obsvr	.033708	.040650	.020000	.019802	.7427
49	Number Drivers in Household	Pre-Tested	2.4324	2.4554	2.2772	2.3590	.4991
		Pre & Post	2.4225	2.3978	2.3288	2.4810	.7918
		Same Obsvr	2.3390	2.4400	2.4151	2.5161	.7848

*Chi-square significance level

TABLE 4.2 (Continued)

Variable Number	Variable Name	Sample	MEANS 8-Hour	4-Hour	Self Teach	WEDMK	ANOVA α
1024	State Record for Two Years: Number Collisions	Pre-Tested	.10465	.15135	.1444	.1734	.3862
		Pre & Post	.13043	.14379	.13740	.17188	.8301
		Same Obsvr	.12766	.16102	.12245	.17895	.6809
1025	State Record for Two Years: Number Single Veh. Coll.	Pre-Tested	.0058140	0.	.0055556	.023121	.0962
		Pre & Post	.0086957	0.	0.	.023438	.0915
		Same Obsvr	.010638	0.	0.	.021053	.2471
1026	State Record for Two Years: Number Injury Collisions	Pre-Tested	.023256	.075676	.055556	.086705	.0796
		Pre & Post	.026087	.078431	.053485	.078125	.2843
		Same Obsvr	.021277	.084746	.040816	.073684	.2141
1027	State Record for Two Years: Number Convictions	Pre-Tested	.37791	.36216	.36667	.33526	.9476
		Pre & Post	.33913	.32680	.37405	.32813	.9371
		Same Obsvr	.35106	.33898	.40816	.34737	.8985
1028	State Record for Two Years: Number Points*	Pre-Tested	.55233	.60000	.58889	.52601	.9495
		Pre & Post	.45217	.50980	.63359	.50000	.7038
		Same Obsvr	.53191	.55085	.66327	.52632	.8773
1074	TARIC Claims for Two Years: Number Collisions	Pre-Tested	.59406	.50485	.66019	.75510	.3857
		Pre & Post	.68852	.56962	.79221	.78378	.5776
		Same Obsvr	.63265	.59322	.77049	.86275	.6098
1075	TARIC Claims for Two Years: Number Liability Collisn	Pre-Tested	.16832	.22330	.31068	.19388	.2758
		Pre & Post	.22951	.24051	.37662	.22973	.3810
		Same Obsvr	.18367	.25424	.34426	.23529	.5540
1076	TARIC Claims for Two Years: Number Comprehensive Clm	Pre-Tested	.28713	.36893	.31068	.44898	.4904
		Pre & Post	.24590	.37975	.29870	.41892	.5769
		Same Obsvr	.24490	.44068	.26230	.33333	.3635
1078	TARIC Claims for Two Years: Number Subrogations	Pre-Tested	.039604	.14563	.077670	.071429	.1263
		Pre & Post	.032787	.16456	.10390	.081081	.1660
		Same Obsvr	.020408	.20339	.098361	.098039	.0911
1077	TARIC Claims for Two Years: $ Cost Damage-Leased Car†	Pre-Tested	96.720	61.634	106.89	84.274	.6561
		Pre & Post	137.53	70.468	124.44	88.549	.6540
		Same Obsvr	135.90	77.356	127.72	90.708	.7393

*Number points are under recorded, as they have been expunged from state records for part of the period measured here.

†Several abnormally expensive claims have been removed to avoid unreasonable inflation of means. Anova significance is based on a log transformation because of skewness.

matched observers did not differ significantly on any of the biographical variables (at the 5% level or better) from the remainder of the matched observer groups. This comparison was carried out with two sample t tests.

Variable Number	Variable Name	MEANS 8-Hour	4-Hour	Self Teach	WEDMK	ANOVA α
10	Salary Grade	14.286	7.1034	8.4483	10.483	†
14	Marital Status	1.1071	1.1724	1.2069	1.1379	.7589
15	Race/Sex	not appropr.	not appropr.	not appropr.	not appropr.	.6876*
18	Year of Birth	25.857	29.345	33.655	29.103	.0155
21	Year Joined Ford	55.464	58.793	59.828	58.172	.2478
32	Commuting Mileage	18.929	26.414	22.103	33.345	.0167
33	Annual Mileage	16.786	16.414	17.724	18.655	.6265
40	H.S. Driver Ed.	0.148	0.207	0.385	0.138	.1055
1024	State Data: Collisions in 2 Years	0.148	0.185	0.185	0.143	.9746
1026	State Data: Injury Coll. in 2 Years	0.037	0.037	0.074	0.071	.885
1027	State Data: Convictions in 2 Years	0.37	0.259	0.296	0.393	.9047

*Chi-sguare significance level
†Assumption of Homogeneity of Variance not met.

TABLE 4.3

Comparison of Selected Biographical and Driving History Variables, by Treatment Group, for Follow-up Tested Participants (N=111)

Sampling assumptions were next examined for the two non-volunteer groups (those who were sent the WEDMK package, and those who were not contacted after the original invitation). These groups are compared in Table 4.4.

TABLE 4.4
Comparison of the Non-Volunteer Treatment and Control Groups

Variable Number	Variable Name	MEANS: WEDMK N=474	MEANS: No Contact After Inv. N=473	Two Sample *t* test α	Chi-square Degree Freedom/ Significance†
10	Salary Grade	11.333	11.330	.9972	
14	Marital	1.2068	1.1691	.1388	
15	Race/Sex	not appropr.	not appropr.	†	5/.8626
18	Birth Year	28.762	29.076	.6602	
21	Year Joined Ford	56.589	57.245	.5407	
		Driver Record Sub-Samples N=54	N=63		
1024	State Record for Two Years: Number Collisions	.12963	.22222	†	2/.4218
1025	State Record for Two Years: Number Single Vehicle Collisions	0	0	-	
1026	State Record for Two Years: Number Injury Collisions	.03704	.03175	.8766	
1027	State Record for Two Years: Number Convictions	.38889	.42857	.7932	
1028	State Record for Two Years: Number Points	.96296	1.1111	.7572	

†Where the assumption of homogeneity of variance is not met (at the 5% level), and for Race/Sex, the table gives the Chi-square degrees of freedom and significance level.

None of the five key biographical variables were significantly different between the two groups. In addition, for subjects for whom driving histories were obtained, two years of state driver record data may be examined in the table; again, the groups do not appear to differ*.

It was also necessary to examine the representativeness of the driver record sub-sample. A comparison of this sub-sample and the remainder of the non-volunteers is given in Table 4.5. The only notable difference is that the sub-sample joined the company, on the average, about two years earlier than the remainder.

TABLE 4.5
Comparison of the Non-Volunteer Driver Record Sub-Sample With the Remaining Non-Volunteers

Variable Number	Variable Name	MEANS: Driver Record Sub-Sample N=117	MEANS: Remaining Non-Volunteers N=830	Two-Sample t test α	Chi-square Degree Freedom/ Significance†
10	Salary Grade	11.692	11.281	†	31/.8709
14	Marital	1.2308	1.1819	.2059	
15	Race/Sex	not appropriate	not appropriate	†	5/.8626
18	Birth Year	27.154	29.167	.0637	
21	Year Joined Ford	55.462	57.276	†	42/.0364

†Variances significantly different at 5% level, but not at 1% level. Also see note, Table 4.4.

*It will be noted that with sample sizes as small as those used here for driver record comparisons, it is sometimes difficult to meet the t test assumption of homogeneity of variances. F tests were performed on the variances; where they were found to differ at the 5% level or better, and also the t statistic was close to the critical value, a Chi-square test was performed. The Chi-square statistic was also used throughout the study for comparisons of race/sex, as it is a strictly nominal variable.

A similar examination of the driver record sub-sample of the 302 cancellations will be found in Table 4.6. Only the length of service with the company appears to differ to any degree, and this is significant at no better than the 9% level.

TABLE 4.6
Comparison of the Cancellation Driver Record Sub-Sample with the Remaining Cancellations

Variable Number	Variable Name	MEANS: Driver Record Sub-Sample N=103	MEANS: Remaining Cancellations N=199	Two-Sample *t* test α	Chi-square Degree Freedom/ Significance†
10	Salary Grade	11.835	10.593	†	21/.8332
14	Marital	1.0583	1.1156	†	1/.1089
15	Race/Sex	not appropriate	not appropriate	†	4/.4040
18	Birth Year	29.262	30.307	.3733	
21	Year Joined Ford	56.233	58.106	.0935	

†Where the assumption of homogeneity of variance is not met (at the 5% level), and for Race/Sex, the table gives the Chi-square degrees of freedom and significance level.

General comparisons between the major groupings of subjects (i.e., volunteers who participated, volunteers who cancelled, and non-volunteers), are a legitimate concern of "immediate" evaluation, and are therefore discussed in Section 5.1.

4.3 VALIDATION OF OPERATIONAL ASSUMPTIONS

As described in Section 2.3, several operational factors were randomized to the extent practical considerations allowed. Several of these were subjected to statistical analysis.

It is particularly important in the evaluation of training

effects to keep the time which elapses from pre-test to post-test constant between treatment groups. The same is true of post-test to follow-up test periods. Table 4.7 shows the means for both periods by treatment group. One way analyses of variance revealed no significant difference between the groups.

TABLE 4.7
Mean Time in Days Between Tests

TEST INTERVAL	8-HOUR	4-HOUR	SELF-TEACH	WEDMK	ANOVA α
Pre-test to Post-test N=556	37.72	36.83	38.38	37.43	.3390
Post-test to Follow-up Test N=113	125.78	121.00	132.9	120.32	.5591

Because of measured differences in observational "style", the allocation of road test observers to the treatment groups potentially could introduce serious bias into the results on the SBTW test. Because the SBTW analyses were carried out exclusively on the 428 cases with matched observers on pre- and post-tests, this same set of cases was used to test the assumption of random observer assignment. The frequencies and Chi-square test details are given in Table 4.8. The significance level of .952 suggests that the four treatment groups are very well matched in this respect.

Of lesser import was the allocation of instructors to the formal training sessions. These were analyzed separately for road and classroom components within the Eight-Hour and Four-Hour courses. A close to random assignment over all training sessions is desireable because of the intentional variability in personal characteristics between the instructors. It is especially desireable for the road training sessions, because therein the

TABLE 4.8
Allocation of Road Test Observers to Treatment Groups
(Frequencies) (Cases with Matched Observers on Pre- and Post-Test Only)

OBSERVER NUMBER	8-HOUR	4-HOUR	SELF-TEACH	WEDMK	TOTALS
1	4	6	3	6	19
2	19	25	24	23	91
3	20	23	23	24	90
4	10	9	12	10	41
5	19	31	23	17	90
6	14	11	11	7	43
7	5	7	6	8	26
8	7	11	4	6	28
Totals	98	123	106	101	428

CHI SQUARE = 11.507 DF = 21 SIGNIF. LEVEL = .952

success of the teaching situation depends heavily on the personal skills of the instructor. It would be particularly disadvantageous to the experimental design if certain instructors had tended to "specialize" in the third and fourth sessions, as this would introduce systematic differences between the Eight-Hour and Four-Hour courses.

Table 4.9 reveals an equitable distribution of road training instructors for the Eight-Hour course. (The differences in the

TABLE 4.9
Eight-Hour Course
Allocation of Instructors to Road Training Sessions (Frequencies)

	INSTRUCTOR NUMBER								INSTRUCTOR UNKNOWN	TOTAL SESSIONS
	1	2	3	4	5	6	7	8		
Road Session 1	8	32	28	9	28	14	5	9	6	139
Road Session 2	5	29	23	16	20	14	6	4	3	120
Road Session 3	2	26	18	16	13	13	4	3	8	103
Road Session 4	4	19	22	14	14	17	6	6	2	104
Total Sessions	19	106	91	55	75	58	21	22	19	

CHI SQUARE = 21.152 DF = 24 SIGNIF. LEVEL = .6063

total number of sessions, whether road or classroom, taught by each instructor, result primarily from different lengths of service.) The Eight-Hour classroom sessions, by comparison, do show some imbalance (Table 4.10). The Chi-square significance of .0002 is perhaps misleading, it is based on rather large marginals which tend to increase significance. The imbalance is not extreme enough to suggest that the four classroom components of the Eight-Hour course were inadequately supplied with the full range of instructors.

As shown in Tables 4.11 and 4.12, both the road and the classroom sessions of the Four-Hour course appear to be satisfactorily distributed among the instructors. Tables 4.10 and 4.12 combined throw additional light on the possibility of unwanted "specialization" in either the first or second pairs of classroom

TABLE 4.10
Eight-Hour Course
Allocation of Instructors to Classroom Training Sessions (Frequencies)

	INSTRUCTOR NUMBER								INSTRUCTOR UNKNOWN	LIBRARY ASSIGNMENT*	TOTAL SESSIONS
	1	2	3	4	5	6	7	8			
Classroom Session 1	13	26	21	25	16	10	10	10	1	6	138
Classroom Session 2	12	10	31	20	14	11	7	9	9	-	112
Classroom Session 3	12	17	14	13	14	20	4	7	2	9	112
Classroom Session 4	8	19	15	6	20	9	10	3	8	8	106
Total Sessions	41	72	81	64	64	50	31	29	20	23	

CHI SQUARE = 60.418 DF = 27 SIGNIFICANCE = .0002

*If only one participant was present, a library assignment was sometimes given in lieu of class.

TABLE 4.11
Four-Hour Course: Allocation of Instructors to Road Training Sessions (Frequencies)

	INSTRUCTOR NUMBER								INSTRUCTOR UNKNOWN	TOTAL SESSIONS
	1	2	3	4	5	6	7	8		
Road Session 1	8	24	36	15	27	19	13	10	4	157
Road Session 2	8	36	27	8	24	16	9	11	7	146
Total Sessions	16	60	63	23	51	35	22	21	11	

CHI SQUARE = 7.52 DF = 8 SIGNIFICANCE = .4817

TABLE 4.12
Four-Hour Course: Allocation of Instructors to Classroom Training Sessions (Frequencies)

	INSTRUCTOR NUMBER								INSTRUCTOR UNKNOWN	LIBRARY ASSIGN-MENT*	TOTAL SESSIONS
	1	2	3	4	5	6	7	8			
Classroom Session 1	19	27	38	13	26	15	6	11	2	3	160
Classroom Session 2	20	24	39	12	21	17	3	8	1	6	151
Total Sessions	39	51	77	25	47	32	9	19	3	9	

CHI SQUARE = 3.461 DF = 9 SIGNIFICANCE = .9432

*If only one participant was present, a library assignment was sometimes given in lieu of class.

sessions. The proportion of sessions 1 and 2 combined, as a percentage of all classroom sessions taught, for each instructor,

ranges from 62% to 81%. For the three instructors who were present for the entire program, and who taught over half of all the sessions, the range was 70% to 81%. This suggests only a moderate amount of "specialization".

The classroom sessions were more standardized than the road sessions, partly because of more intensive instructor preparation for the former, and partly because of the use of significant amounts of "packaged" audio-visual material, including some programmed learning. Therefore, it is reasonable to accord these variations in classroom instructor allocation only minor significance.

Finally, although there were some systematic differences in the mean time of day between treatment groups, this was found statistically to have negligible effect on key test scores.

5.0 DISCUSSION OF RESULTS

The results of all data available at the time this report was prepared are discussed in three parts: immediate evaluation, based on program flow characteristics and much of the survey data; intermediate evaluation, which covers the test data; and comments on the prospects for ultimate evaluation.

5.1 IMMEDIATE EVALUATION

Two kinds of data are available to measure the immediate impact of the program. The first is obtained by monitoring "flow" records of volunteering, cancellation, and attendance characteristics. The second is provided by survey data.

Flow Characteristics

The basic information in flow has already been presented in Table 4.1. A little over half of the original 2,000 invitees accepted, but each of the four equally-sized volunteer treatment groups lost about 30% of their number as cancellations In many cases, this was the result of irreconcilable time conflicts.

It is particularly important in understanding the feasibility of this kind of program to compare the characteristics of those who volunteered, with those who did not. A further comparison can be made between volunteers who actually participated and those who cancelled.

All three groups are compared on two biographical variables of outstanding importance, age and salary grade, in Figures 5.1 and 5.2. These variables are, of course, interrelated, and to some extent they show similar trends. Those who originally volunteered tend to overrepresent the middle salary grade and age ranges. This is particularly true of salary grade, which shows substantially higher variance for

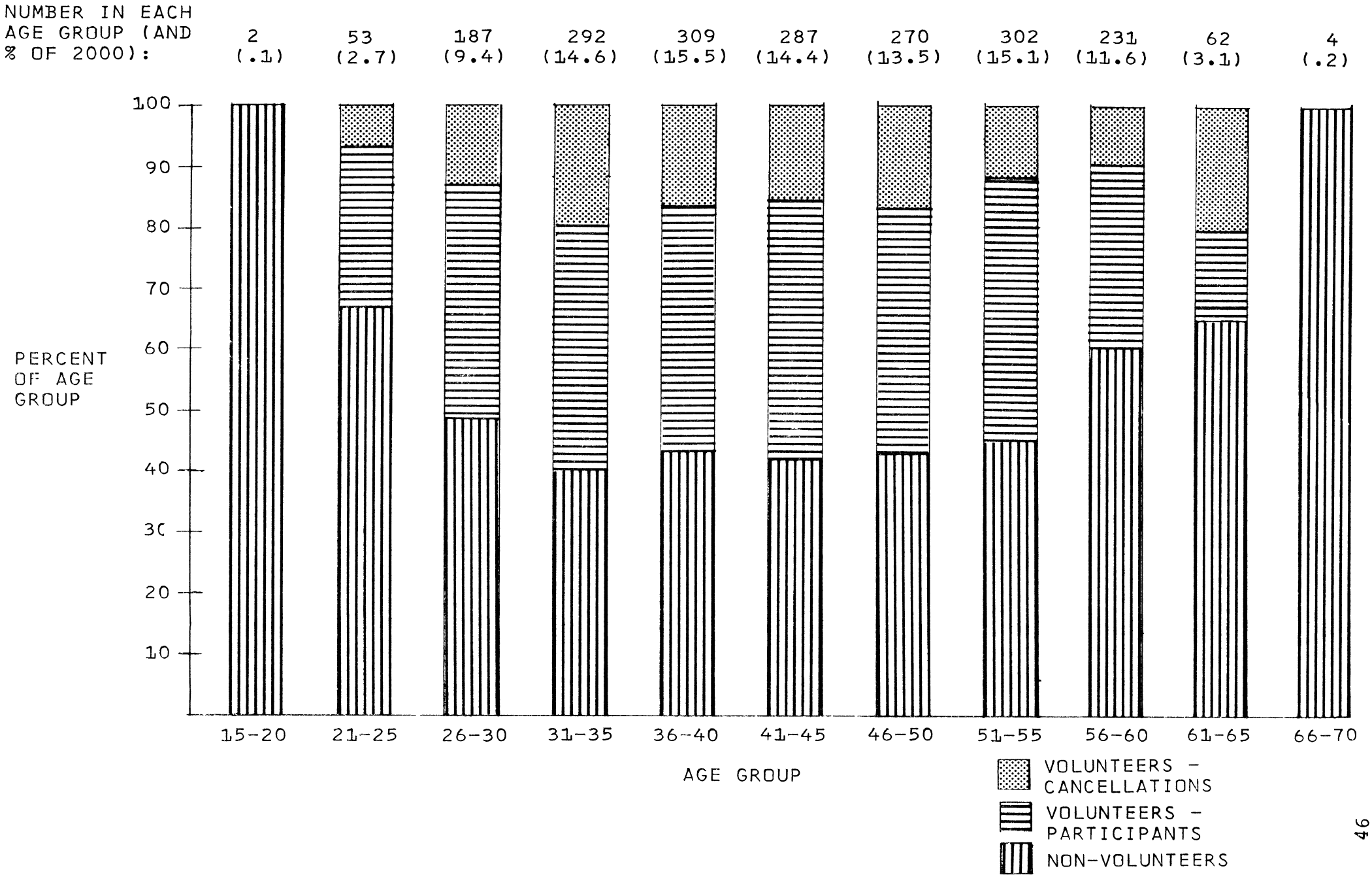


FIGURE 5.1 Response to Program Invitation by Age Group

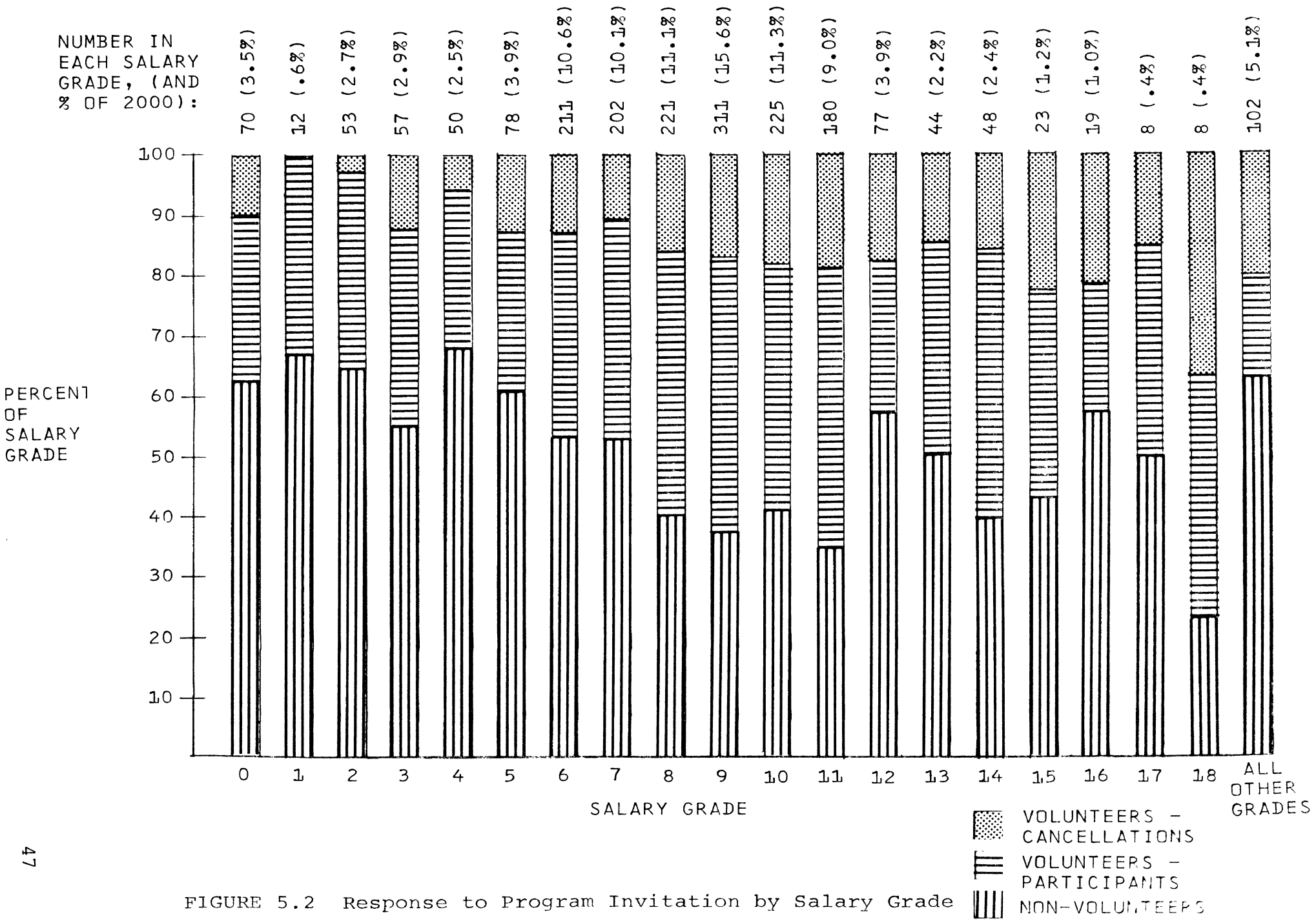


FIGURE 5.2 Response to Program Invitation by Salary Grade

non-volunteers than for volunteers. However, there also is some tendency for the middle salary grade and age ranges to be over-represented in the cancellations, with the result that the participant volunteers were less extreme in this respect. A relatively high proportion of the thinly-represented top four salary grades fall into the cancellation category. Overall, the stronger appeal of the program to the "middling" ranges could result partly from the desire of those in promotion positions to cooperate with company programs.

A consistent weakness of driver programs offered to volunteers is the lack of information to support their representativeness of larger driving populations. *Ex post facto* studies of some high school programs suggest that those who volunteer may have important differences in social and driving history characteristics compared to those who do not. Table 5.1 provides a comparison on key biographical and driving record variables between all three groups under consideration. Means are given for all variables (except Race/Sex), to assist in interpreting the meaning of various statistically significant differences found between the groups. Salary grade appears to be the only variable for which sizable group differences exist in the means. Between group variances, as already noted, are far from homogenous, making it difficult to test means directly. However, given these samples, a Bayesian Posterior Probability of .9993 was calculated that volunteer participant mean salary grade is lower than that of the non-volunteers. As Table 5.1 shows, it is reasonable to infer from Chi-square statistics that, both volunteer groups differ very significantly from the non-volunteers on salary grade. A partial explanation for this difference may be that the more senior employes tend not to volunteer for programs which commit their time during working hours.

There is a slight tendency for fewer volunteers to be married, despite that they also tend to be slightly older than the volunteer groups. There is also a tendency for volunteers

TABLE 5.1 Overall Comparison of Volunteer Participants, Volunteer Cancellations, and Non-Volunteers

Variable Number	Variable Name	MEANS			TWO-SAMPLE *t* SIGNIFICANCE, OR *CHI-SQUARE SIGNIFICANCE IF VARIANCES NOT HOMOGENOUS*		
		Volunteered & Participated N=750	Volunteered & Cancelled N=302	Non-Volunteers N=947	Vol. Partic. vs Vol. Cancel	Vol. Partic. vs Non-Vol	Vol. Cancel vs Non-Vol
10	Salary Grade	9.4747	11.017	11.332	*.1307*	*.0000*	*.0000*
14	Marital	1.1427	1.0960	1.1880	*.0414*	*.0132*	*.0002*
15	Race/Sex	--------NOT APPROPRIATE--------			*.5462*	*.0474*	*.0173*
18	Birth Year	29.755	29.960	28.919	.7643	*.0232*	*.0040*
21	Year Joined Ford	58.047	57.467	57.052	*.2084*	*.0256*	*.3989*
1024	State Record for Two Years: Number Collisions	.14366	.20388	.17949	*.2704*	.3439	.6656
1025	State Record for Two Years: Num-Single Vehicle Collisions	.00845	0.	0.	*.8769*	*.9960*	--
1026	State Record for Two Years: Number Injury Coll.	.06056	.04854	.03419	*.2905*	*.5338*	*.2877*
1027	State Record for Two Years: Number Convictions	.36056	.48544	.41026	*.4268*	*.5457*	.5059
1028	State Record for Two Years: Number of Points	.56761	.75728	.73504	*.2896*	*.7452*	.9184

to include slightly fewer women and non-whites, but the overall numbers of these are not large. There seems to be little difference in the average length of service with the company.

Comparison of all State driver record variables available reveal no significant differences between the three groups. The higher mean collision rate of the cancellation group, although not statistically significant, may support some anecdotal evidence that recent accident experience can have an inhibiting effect on voluntary participation.

In general, there is no reason to suppose that any one of the personal characteristics has much relevance to the fact of cancellation. However, neither is there substantial evidence that the results of this program could not be extrapolated to the large majority of salaried employes in corporations such as Ford Motor Company. The extent to which results may be generalized to the wider driving public is not known at this time.

Another aspect of flow which is of considerable interest is that 190 of the employes who participated in at least a pre-test dropped out of the program before their post-test. These comprized 33% of the Eight-Hour group, 18% of the Four-Hour group, 25% of the Self-Teach group, and 27% of the WEDMK group. The difference in dropout rate between the Eight-Hour and Four-Hour courses almost certainly reflects the greater convenience of the latter treatment. The remaining two groups were only required to attend the pre-test and post-test, and it is reasonable to expect their dropout rates to be higher once they realized that, other than individual study, their participation was limited to these sessions.

A comparison was made of available biographical characteristics between dropouts and those who completed the program. This included the information gathered in the questionnaire at pre-test. A number of significant factors were revealed. Dropouts, like cancellations, tended to overrepresent the middle range of salary grades. They also tended to drive company

cars more often, drive further to work, and have a higher annual mileage, than those who stayed in the program. Other variables, including State driver record data, were not significantly different.

The point at which dropping-out occurred for those who were scheduled for instructional sessions was also examined. Table 5.2 shows that about one-half of the dropouts in both formally trained groups did not begin the training sessions. Thereafter, approximately equal numbers of participants dropped out after each session, except for the last session of the Eight-Hour course. For comparison, attendence is also shown for those pre- and post-tested. More participants would have dropped out after completing all of their assigned instruction but for a special effort to re-schedule those who had missed their post-test.

TABLE 5.2
Number of Instructional Sessions Completed

		0 Tested Only	1	2	3	4	Group Totals
8-Hour	Dropouts	29	11	10	8	2	60
	Pre & Post	4	1	12	27	76	120
4-Hour	Dropouts	15	9	8	-	-	32
	Pre & Post	8	23	148	-	-	169

The pre-test results of dropouts were also analyzed to examine the possibility that they were more or less capable in some aspects of driving ability than those completing the course. Differences are close to zero for the Subjective Behind-The-Wheel scores and for the Perception of Hazards Test. The dropouts do seem to be slightly higher on the Unusual Uses Test, possibly indicating a greater familiarity with problem-solving approaches.

When participants commented about the need to drop out, this was recorded. By far the most frequent reason given was

workload. Other factors may have included company employes' growing general knowledge of the nature of the various treatments in the program. Several press releases were made within and outside the company about the four different courses. This may have led more Self-Teach and WEDMK participants to feel "short-changed" later in the program, and indeed the dropout rates tended to increase in the later cycles. Very few dropouts at any time were openly hostile to the program as they had found it, and many would have completed all their sessions under more flexible scheduling arrangements than were possible in these circumstances.

The final flow characteristic to be considered concerns one of the base conditions in the experimental design, namely equal access to the program's informational content through the Audio-Visual library. Records of library usage do not fully reflect its use at all times throughout the program, but there is no reason to suspect that they do not reveal the pattern of usage by each treatment group. Table 5.3 shows that, as expected, a much smaller number of participants in the Eight-Hour and Four-Hour groups made use of the facility. These two groups also used fewer items. The number of items used refers to self-instruction audio-visual modules and to take-home booklets.

TABLE 5.3
Use of Library Materials, by Treatment Group

	8-HOUR	4-HOUR	SELF-TEACH	WEDMK	TOTAL
Number of Participants Using Library	16	19	30	66	131
Average Number of Library Items Used	3.62	3.05	6.2	5.48	5.81

It should be noted that the Self-Teach package included all of the booklets which cover the content of the program. Library usage was not much of a substitute for dropping classes; only 11% of the dropouts used it, compared to 20% of those completing the program.

Various analyses were performed to compare those who used one to four library items with those who used five or more, and also with those not using the library at all. The library was used more extensively and by more participants in the later months of the program. As Table 5.3 suggests, the majority of "five plus" users were in the Self-Teach and WEDMK groups. They tended to be younger, and to drive more miles per year than the non-users. The "one to four" group by contrast, on the average, were slightly older and drove fewer miles per year than the non-users. Other biographical characteristics, not directly a function of age, were not significantly different.

Survey Data

Approximately 400 surveys were received back in time for the analyses discussed in this report*.

The survey employed a number of ordinal-type multiple choice questions, responses to which are summarized in Table 5.4. The two questions on general approval for the program, out of its usefulness, or as company policy, do not distinquish markedly between the treatments. The same is true of a similar "company policy" question which was answered on a yes/no basis in the pre-test questionnaire; in this case, the favorable vote was almost unanimous. However, the usefulness question is more favorably answered, the greater the intensity of treatment. All four groups appear to have received the message, contained in the program material, that nine out of ten

*The data file will be updated later with all additional responses and made available for further analysis.

TABLE 5.4
Responses to Ordinal-Type Questions on Mailed Survey,
By Treatment Group - Average N = 396

Variable Number	Variable Name		Eight-Hour	Four-Hour	Self-Teach	WEDMK
			MEANS			
244	How Useful Was Program (1 High to 5 Low)		2.53	2.69	2.90	3.07
253	Approve Company Giving Program (1 Low to 5 High)		4.34	4.07	4.16	4.13
256	Self Rate Driving (1 High to 5 Low)	-Before Program	2.45	2.35	2.23	2.11
257		-After Program	2.16	2.04	2.07	2.01
			PERCENTAGE OF RESPONSES			
262	Would you like to retake such a program...	...Every 2 yrs.	26%	12%	25%	14%
		...Every 3 yrs.	21%	34%	22%	29%
		...Every 4 yrs.	15%	15%	16%	25%
		...Not again	19%	27%	28%	23%
		...Other	19%	12%	9%	9%
264	Would you like the program offered to your family...	...Yes	87%	81%	84%	78%

drivers believe they are above average; in each group, self-ratings of driving ability were lower after the program than before, even though both ratings were requested on the same survey in adjacent questions. The Four-Hour group show the greatest drop in self-ratings.

Treatment group seems to have had little effect on how often a participant felt it would be helpful to take a program such as they underwent, or upon whether participants would like it offered to their families. The Four-Hour group is perhaps a little more unfavorable on the latter question

than might be expected from the popularity of that treatment. Responses to a specific question relating to participants' beliefs about who has the most accidents are summarized in Table 5.5. The differences noted are not significant statistically, but the trend is favorable to the program; the greater the amount of training, the greater the recognition that the average driver, and not popular scapegoat groups, is responsible for the great majority of accidents.

TABLE 5.5
Percentage Responses, by Treatment Group, to Survey Question "Who Has the Highest Percentage of All Traffic Accidents?"

Response	Eight-Hour N=87	Four-Hour N=103	Self-Teach N=101	WEDMK N=93
Drivers Who Have Already Had a Serious Accident	2.3%	4.9%	5.9%	7.5%
Drivers With Grossly Anti-Social Behavior	25.3%	28.2%	32.7%	38.7%
"Average" Drivers	72.4%	66.0%	60.4%	53.8%
	100%	100%	100%	100%

CHI SQUARE = 8.225 DF=6 SIGNIFICANCE = .2221

The remainder of the survey consisted of open-ended questions about all aspects of the program. It is worthwhile to examine responses to two types of question.

The first type asks only for a statement of what aspects were most or least useful. Table 5.6 lists the aspects which were most frequently mentioned, and then notes the percentage of responses in each treatment group which named these aspects, either as most useful or least useful. The columns do not sum to 100% because a large number of infrequent responses have been excluded. In general, road instruction and audio-visual

TABLE 5.6
Responses to Survey Questions on the Most and Least Useful Aspects of the Program, by Treatment Group

ASPECT	Percent of Respondents in Each Treatment Group Naming Aspect of Program as Most or Least Helpful (Rounded)									
	8-HOUR N=92		4-HOUR N=102		SELF-TEACH N=100		WEDMK N=93		TOTAL N=387	
	Best	Least	Best	Least	Best	Least	Best	Least	Best	Least
Classroom Instruction	8%	5%	7%	1%	-	-	1%	1%	4%	2%
Classroom Tests	-	3%	-	2%	4%	2%	3%	2%	2%	13%
Road Instruction	24%	8.5%	20%	10%	3%	3%	1%	1%	12%	6%
Road Tests	6.5%	5%	5%	2%	7%	10%	11%	35%	8%	3%
Perception of Hazards Test	2%	2%	2%	9%	6%	12%	9%	3%	5%	7%
*Audio-Visual Aids	5%	7.5%	10%	10%	4%	6.5%	20%	4%	10%	7%
Feedback From Tests	1%	1%	1%	3%	-	13%	-	12%	0.6%	7%
Tests in General	-	4%	-	8%	-	4%	1%	1%	0.2%	5%
"None" or not Applicable	4%	22%	7%	28%	6%	17%	9%	12%	7%	21%

NOTE: In response to a request for specific examples of how the program has helped, 12 subjects claimed they had actually avoided accidents because of their training, and 23 said that they were aware of specific danger situations because of, and since, the program.

*Of 35 subjects giving a second response to the question "What was most useful", 10 named Audio-Visual aids.

aids are the most commended. Testing comes in for criticism in various ways, primarily because of frustration with the rule that they could not be told the results (necessary to prevent test information from being exchanged by participants). Overall, about one-fifth of the participants made the generally favorable comment that "nothing" was least useful.

Within and between the treatment groups, there are some other interesting inferences to be made. A significant number of the Self-Teach and WEDMK participants regarded both the classroom and the road tests as training, especially the latter. Road instruction was the most favored aspect overall, but it should be noted that in both the Eight-Hour and Four-Hour groups, for every two participants considering it most useful, there was about one who considered it least useful. Interestingly, a noticeable number of those disliking road instruction of testing mentioned the Perception of Hazards tests as most useful. Yet overall, the P.O.H. test was rather controversial with the participants; antipathy to it is probably concealed in one of two of the other aspects, especially "classroom tests", and it possibly increased disapproval of visual aids. The strong WEDMK vote for visual aids may reflect library usage; this was certainly the most interesting instructional medium available to them. Their marked disapproval of road testing is hard to explain.

The overall patterns of best-least responses are somewhat similar for the Eight-Hour and Four-Hour groups, although the latter seems to have greater internal disagreement and more general satisfaction. With this set of aspects, the remaining two groups can really only comment on the tests, which on balance they did not favor.

The second type of open-ended question, despite appearances, is importantly different from that just discussed. It concerns suggested changes and additions to the program, as prompted by the four row headings shown in Table 5.7: content,

TABLE 5.7
Responses to Survey Questions on Changes and Additions Desired for the Program, by Treatment Group

	8-HOUR			4-HOUR		
	Number Resp.	Unfavorable	Favorable	Number Resp.	Unfavorable	Favorable
Program Content	70	6%	34%	86	1%	37%
Instructors	61	10%	8%	72	13%	13%
Teaching Methods	66	12%	21%	73	18%	10%
Instructional Materials	65	20%	22%	76	21%	13%

	SELF-TEACH			WEDMK		
	Number Resp.	Unfavorable	Favorable	Number Resp.	Unfavorable	Favorable
Program Content	81	9%	32%	84	4%	37%
Instructors	67	10%	15%	69	4%	9%
Teaching Methods	75	23%	27%	77	3%	13%
Instructional Materials	68	16%	12%	70	13%	14%

instructors, teaching methods, and instructional materials. This table summarizes the percentage of favorable and unfavorable responses (within each treatment group), under each of the four headings separately. "Unfavorable" in this context means

essentially that the program contained too much of some named feature of content, methods, instructors or materials. Favorable means that more is desired. The purpose of Table 5.7 is to compare the percentage of respondents who were thinking about changes in negative terms, with the percentage of those considering the reverse. Each percentage relates to the number of respondents shown in each row. If the percentages of *neutral* responses were also shown, each row would sum to 100%.

About one-third of participants in all treatment groups desired extension of the content of the program. (Many responses have to do with more road work and critical or emergency driving situations.) Interestingly, the Self-Teach group is the most desirous of reductions in content.

Comments on instructors refer to their role as testers/observers for the Self-Teach and WEDMK groups. Most responses about instructors were relatively neutral, with no strong trends towards increasing or decreasing their skills and availability. The Self-Teach and WEDMK groups are high on "favorable" because a significant number simply stated their desire to be formally instructed. A similar difficulty exists in interpreting the methods figures; the last two groups say a fair amount against self-instruction or for formal instruction. But the response patterns for the Eight-Hour and Four-Hour groups suggest that with greater exposure to training, participants may well demand more variety of learning situations.

There was considerable comment on instructional materials with the Eight-Hour people again having the most suggestions for additions; however, in both formally training groups, over 20% would eliminate some of the current materials. As with the best/least responses, the Self-Teach and WEDMK groups tend to interpret the test situation in instructional terms, and most of their "unfavorable" comments refer to testing materials.*

---

*A very detailed breakdown of these, and all other open-ended responses, is given for each treatment group in Volumes 4B to 4E (Volunteer Participant File Means and Marginals); these should be used with Volume 3 (Code Book for Volunteer Participant File).

Some additional survey type information is provided by responses to the request for near miss counts. At post-testing and follow-up testing, participants were asked to transfer this information from the record cards they were given at pre-test. Because of obvious biases and ambiguities in keeping such records, it is unwise to draw too many conclusions from near miss rates. However, the percentage of each treatment group taking the trouble to keep a record is itself an index of interest in, and cooperation with, the program. In some cases, a high rate may also indicate interest in this as an educational exercise. This data is given in Table 5.8.

TABLE 5.8
Near Miss Data: Percentage of Each Treatment Group Responding, and Reported Near Miss Rates

	Post-Test			Follow-up Test		
	N	% Resp =	Average=	N	% Resp =	Average=
8-Hour	54	45.0	.25	9	33.3	.52
4-Hour	70	43.4	.45	11	37.9	.23
Self-Teach	47	32.6	.30	7	24.1	.80
WEDMK	40	30.0	.23	3	10.3	.75

The Eight-Hour and Four-Hour respondents have clearly been the most responsive.

Finally, a single attempt was made to obtain non-volunteers' opinions on the program, and their reasons for not accepting the invitation. Approximately 350 randomly selected non-volunteers were mailed a two question survey (a copy of which is included as Appendix 8) in July or November 1972. Sixty-six replies were received, of which eleven expressed some disapproval of the program. Most of the remainder thought that the company should offer advanced driving instruction on the grounds that it was a logical activity for an automobile company to pioneer (for safety or for company image), a desirable job benefit, or as an aspect of "product responsibility". Forty-seven, including some

of those not in favor of the program, described workload or travel as the reason for not participating; nine wanted to participate but had not successfully transmitted their acceptance of the invitation; and three asked if they could reverse their refusals.

5.2 INTERMEDIATE EVALUATION

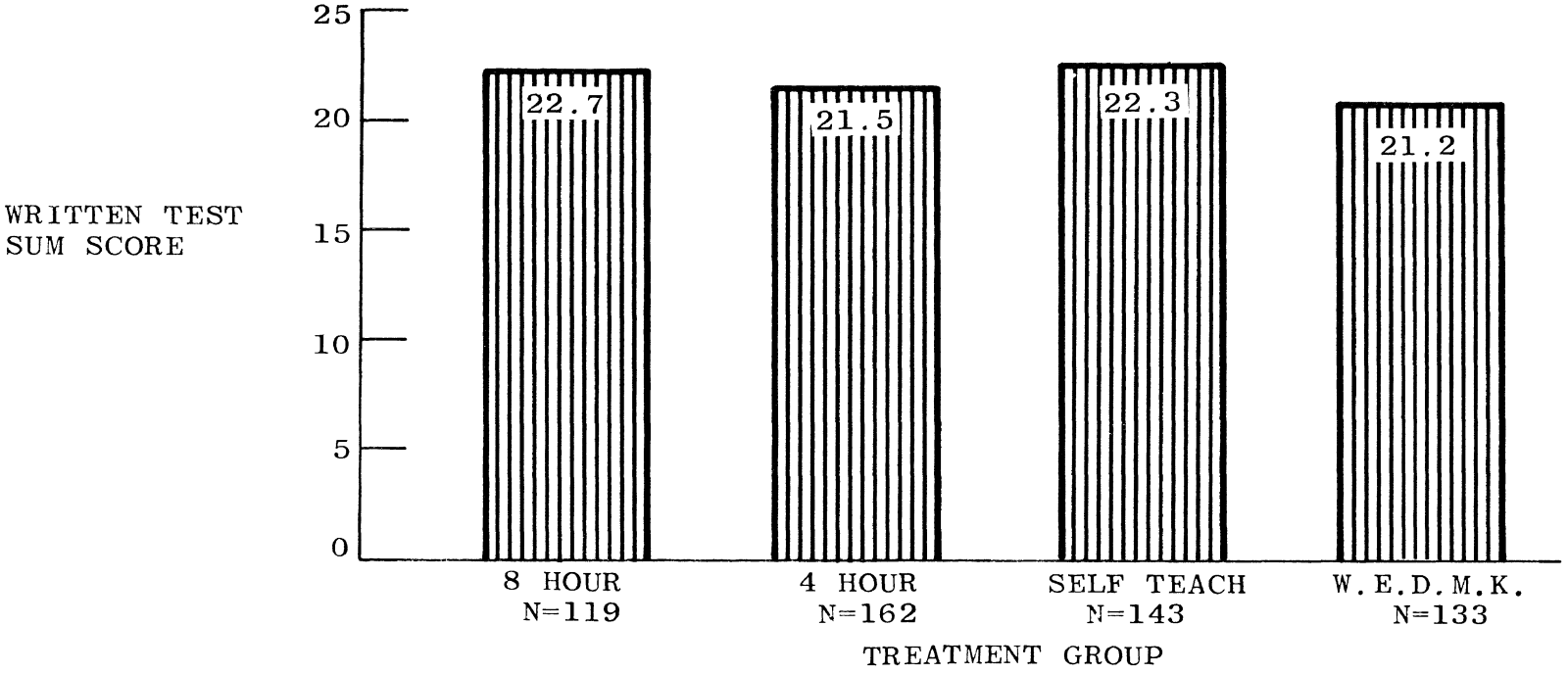
The results from the various tests are discussed in the order suggested by the evaluation model.

Written Test (of Content Acquisition and Knowledge of Driving Techniques and Problems).

Content acquisition comprised Parts II and III of the test, which was not given at pre-testing. Part II, contained more specialized terminology from the program content, and as shown in the summary table below Figure 5.3 this sub-score did tend to reflect the intensity of training at post-testing. Part III contained questions on Strategic Positioning. The Eight-Hour group had the highest mean score, but overall there was very little difference between the groups, each averaging about five questions answered correctly out of seven. By follow-up testing, the Eight-Hour group was approximately half a question ahead of all three remaining groups on Part II. Follow-up test scores means on Part III remained almost unchanged.

On both of the Knowledge portions of the test (Parts I and IV), the Eight-Hour and Self-Teach group means were slightly higher at post-testing, as shown in Figure 5.3. This held true at follow-up testing for Part I (General Driving Knowledge), which showed a decrease of about one quarter of a question for the Eight-Hour and Four-Hour groups, and of about one-half a question for the WEDMK group; the Self-Teach mean score was unchanged. On Part IV (Emergencies and Other Driving Problems), the pattern of changes was reversed, with a decline of about one-third of a question for all groups except Self-Teach, which improved very slightly; this left the Eight-Hour group about

Parts II and III test Content Acquisition
Parts I and IV test Driving Knowledge



SUMMARY OF SUB AND SUM SCORES:

PART	MAX. POSS.	8-HR	4-HR	S.T.	WEDMK	ANOVA F TEST SIGNIF
I	10	6.9	6.7	7.3	6.8	.0071
II	6	4.7	4.5	4.4	4.1	.0002
III	7	5.2	4.7	4.8	5.0	.0027
IV	7	5.9	5.5	5.8	5.3	.0021
SUM	30	22.7	21.5	22.3	21.2	.0003

FIGURE 5.3 Written Test Sum- and Sub-scores at Post-testing

half a question ahead of each of the others, whose mean scores were now very similar. The sub scores for all four parts of the Written Test are shown at post-testing in Figure 5.3. The Eight-Hour and Self-Teach groups were generally better than the Four-Hour and WEDMK, and this pattern remained at follow-up testing after some minor fluctuations in group means. It will be noted that ANOVA *F* test significance levels for between-group differences are all high at post-testing, although this results partly from group sizes as large as 120 to 160. At follow-up testing, with group sizes of 28 to 30, the sum score significance level was about 7%, and Parts I and II differed between groups at better than this level. The size of the differences in real terms are quite small, however. For example, the difference between maximum and minimum mean sum scores is only 1 1/2 questions out of a possible maximum of 30. Nevertheless, a number of conclusions may be drawn.

Up to four months following course completion, the tendency for the more intensively trained groups to be ahead only where special terminology was used suggests that formal training has a cognitive effect, but that the majority of the information contained within the program may, as expected, be provided in a number of equally successful ways.

The predominance of the Eight-Hour and Self-Teach groups on the knowledge scores is not easy to explain. One possible reason, which could also apply to the pattern of sum scores, is that the Self-Teach package was about as successful in transferring general driving information as the Eight-Hour course; the Four-Hour group could be lower than either because they were exposed to a more limited range of teaching materials.

Part I questioned general knowledge less closely linked to the program content than Part IV; that the Self-Teach group remained ahead only on Part I at follow-up testing offers some support for formal training as a way of transferring new knowledge. Against all these comments, however, it should once more be noted that the differences are not large.

A special analysis was made of the Written Test scores of those who used the library. No strongly significant differences between users and non-users were found. The strongest result was a higher Part II sub score for WEDMK library users compared to the remainder of the treatment group; this was significant at the 12% level.

Finally, although group mean scores reflect on the average, 73% success on the Written Test, the poorest sub scores are found on Part I, General Driving Knowledge. This was despite the fact that all participants were given the "What Every Driver Must Know" booklet. This implies that adult drivers have some relatively commonplace information needs which are not met by conventional sources. That the treatments were close to equally successful in transferring the more specialized information in the program does not imply that they were equally successful in encouraging its application. It is thus necessary to test more complex levels of behavior.

Perception of Hazards Test

As described in Section 3.2, it was necessary to use two versions of this test on a random split-half basis in order to examine pre-test/post-test differences. Because the two versions were not pre-tested for comparability, difference scores were examined separately according to the order in which the versions were taken. It was apparent that the versions were not comparable in difficulty, because difference scores were negative for one version order, and positive for the other. To take account of this, version order was used as one factor in a two-way univariate Analysis of Variance model. The unweighted means technique (as described by Scheffe, 1959[11]) was used on the post-test/pre-test difference scores. Table 5.9 summarizes the results from this analysis. As expected, the order in which tests were taken was highly significant ($\alpha = <.001$). Neither treatment nor the interaction of treatment and test order were significant, although the BA test application did reflect the intensity of

TABLE 5.9
Perception of Hazards: Two Way ANOVA of
Post-Test Minus Pre-Test Difference Scores

Difference Scores (and Cell Sizes):

	8-Hour	4-Hour	Self-Teach	WEDMK	Sum
Test Order AB	-1.515 (68)	-.716 (81)	-.274 (84)	-.843 (70)	-3.347
Test Order BA	1.082 (49)	.857 (77)	.577 (52)	.441 (59)	2.956
Sum	-.433	.141	.303	-.402	-.391

GRAND MEAN = -.042

Deviation From Grand Mean:

Row Effects (Order):

Order AB: -.788
Order BA: +.788

Column Effects (Treatment):

8-Hour: -.167
4-Hour: +.120
Self-Teach: +.201
WEDMK: -.152

F Statistic	D.F.	Significance, α =
Row Effects (order): 31.22	1,65	<.001
Column Effects (treatment: .96	3.65	> .25
Interaction: 1.25	3,65	> .25

of treatment. Not only are the treatment effects negligible, but the differences between post- and pre-test are very small.

At follow-up test, participants re-took the version they were given at post-test. Descriptive analyses of follow-up test scores and follow-up/post-test difference scores, by test version, revealed small and erratic differences between group means. On these results, and because of the smallness of the follow-up samples after splitting, no further analysis was attempted.

The inability of this test to distinguish between treatment groups must be interpreted with caution. Although it is possible that none of the treatments affected the ability of participants to detect critical hazard information, the validity of the test as applied must be questioned. It was of poor technical quality, especially relative to other audio-visual materials used in the program, and some commented that this affected the willingness of participants to make the effort necessary to respond to the test situations. The average score was approximately 15 out of a maximum of 23; this is high enough to suggest that few items were left unanswered, but it is not possible to know the effect of motivation to complete the test. The rationale for a test of this kind of perceptual ability remains strong; this and related techniques deserve further development.

Unusual Uses Test

This test was also assigned on a split-half basis, and with respect to version difficulty, it suffered from the same problem as the previous test. Therefore, the two way analysis of variance technique with unweighted means was again used. The results of the analysis are summarized in Table 5.10. Test version order is highly significant. The treatment effect is also shown in Figure 5.4, and is significant at the 10.5% level. The Eight-Hour and Four-Hour groups performed about twice as well as the Self-Teach group, although their improvement only amounted to about one-fifth of a response over pre-test scores averaging approximately seven.

The interaction between treatment and test order was also significant at better than the 3% level. This presents some difficulty of interpretation; it implies that one or more treatment groups were especially sensitive to certain test version orders. It should also be considered in relation to the marked drop in WEDMK scores between pre- and post-testing; this could reflect a lower level of interest in, and desire to cooperate with, the post-test.

TABLE 5.10
Unusual Uses: Two Way ANOVA of Post-Test Minus Pre-Test Scores

Difference Scores (and Cell Sizes):

	8-Hour	4-Hour	Self-Teach	WEDMK	Sum
Test Order AB	-1.074 (68)	-1.025 (81)	-1.291 (86)	-.930 (71)	-4.320
Test Order BA	1.857 (49)	1.844 (77)	1.864 (52)	.283 (53)	5.848
Sum	.783	.819	.573	-.647	1.528

GRAND MEAN = -.189

Deviation From Grand Mean:

Row Effects (Order):
Order AB: +1.269
Order BA: -1.269

Column Effects (Treatment):
8-Hour: .206
4-Hour: .216
Self-Teach: .091
WEDMK: -.514

F Statistic	D.F.	Significance, α =
Row Effects (order): 113.79	1,64	<.0001
Column Effects (treatment): 2.12	3,64	.105
Interaction: 3.52	3,64	<.025

Exploratory analyses of the follow-up test results include one way analysis of variance of the raw scores, taking the two test versions separately. As with the previous test, the split-halving of the 116 follow-up test subjects resulted in very small treatment group cells. There were no significant differences between treatment groups on either version, and scores

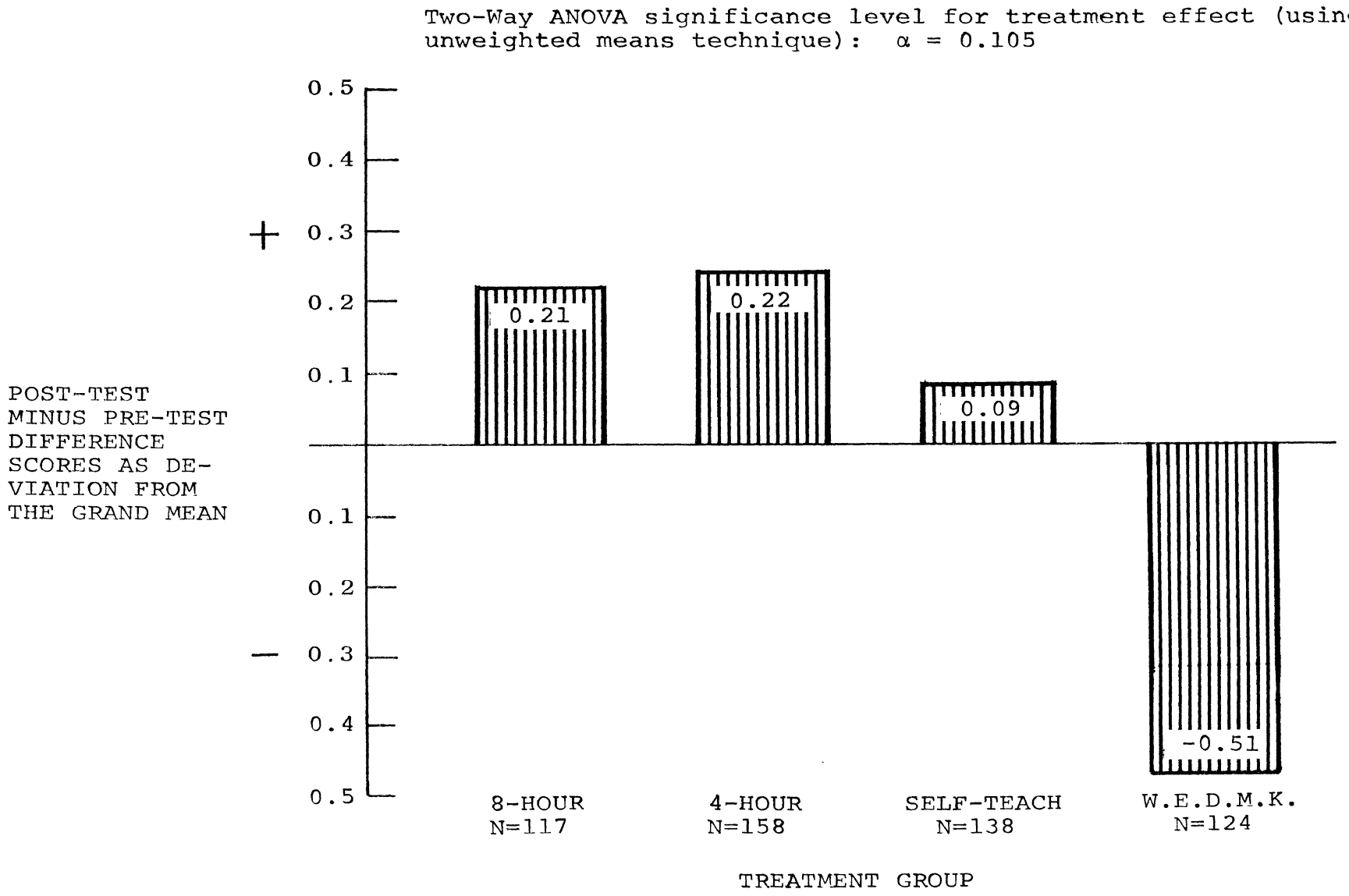


FIGURE 5.4 Unusual Uses: Post-Test Minus Pre-Test Difference Scores, as Deviation From the Grand Mean for Both Test Version Orders

were very close. However, there is some indication of a slight increase in performance by all except the Four-Hour group for those who took the easier version at post-test and follow-up test. The results for those who took the more difficult version on both these occasions are erratic. Overall, the follow-up test did not provide sufficient information to determine whether or not the effects recorded at post-test were retained.

This application of the Unusual Uses technique was a necessary speculative attempt to provide a simple measure of the ability of drivers to think imaginatively about difficult driving situations. In the absence of more sophisticated instrumentation for "process" sub tasks, this provided some evidence that the Eight-Hour and Four-Hour courses raised the ability of drivers in this respect.

Objective Behind the Wheel Test

Analyses were performed with respect to two questions. Firstly: are there differences between treatment groups, which may be detected by the objective measures? Secondly: to what extent can the objective measures differentiate between high scorers and low scorers on the Subjective Behind the Wheel Test? Descriptive analyses led to more detailed examination of these questions for four major types of data: Total Time elapsed during test, Steering Reversal Rates (per minute), Brake Applications (sum score), and Index Numbers (a composite of speed and steering reversals). The results of all these analyses are summarized in Table 5.11. This table also includes some recommendations for additional analysis techniques which might be used in this context for some of the data.

Total time is of interest because unusually slow or fast drivers are potentially more hazardous than those who maintain near-average speeds. Overall, the participants varied less at post-test than at pre-test, as shown by a lower ANOVA *F* statistic. A more accurate measure of this is given by use of

TABLE 5.11

Summary of Results of Analyses of Objective B.T.W. Data

<table>
<tr><th>DATA TYPE</th><th>DIFFERENCE BETWEEN TREATMENT GROUPS</th><th>COMPARISON WITH SUBJECTIVE BEHIND-THE WHEEL DATA</th></tr>
<tr><td>TOTAL TIME</td><td>1. ANOVA OF RAW SCORES: Significant at α = <.001 at pre- and post-test, but overall variance lower at post-test (F statistic dropped from 14.5 to 6.4).
2. ANOVA OF ABSOLUTE VALUES OF Z SCORES: Significance levels: pre-test, α = .0001; post-test, α = .6074.
3. ANOVA OF DIFFERENCE POST PRE-TEST ON ABSOLUTE VALUES OF Z SCORES: Significance: α = .3951.
ADDITIONAL PROCEDURES RECOMMENDED: 2-way ANOVA on treatment and pre-post.</td><td>PRODUCT-MOMENT CORRELATIONS OF ABSOLUTE VALUE OF Z SCORES WITH SELECTED SBTW SCORES: No significant correlations obtained at α = <.05, with SBTW Sum Scores (including sub-score consisting of freeway items only) or average headway: Correlation of -.2037 with average headway on follow-up test (α = <.1).
ADDITIONAL PROCEDURES RECOMMENDED: Rank correlations.</td></tr>
<tr><td>STEERING REVERSAL RATE</td><td>1. ANOVA OF RAW SCORES FOR SIX SEQUENCES OF TEST ROUTE, THE AVERAGE OF THESE, AND RATIO OF FREEWAY STRESS/PRE-STRESS AT PRE-, POST-, AND FOLLOW-UP TESTING: No significant differences at α = <.05 except for recovery-left turn sequence at pre-test.
2. ANOVA OF DIFFERENCES BETWEEN POST-TEST AND PRE-TEST RAW SCORES: No significant differences at α = <.05.</td><td>PRODUCT-MOMENT CORRELATIONS FOR PRE-TEST POST-TEST AND POST-PRE DIFFERENCES ON: AVERAGE RAW SCORES AND FREEWAY STRESS/PRE-STRESS RATIOS; CORRELATED WITH SBTW SUM SCORES (MAX=21) AND AVERAGE HEADWAY FOR PRE-TEST, POST-TEST AND POST-PRE DIFFERENCES: Correlations significant at the α = <.05 level:
Pre-test average SRR, with headway difference (.1346).
Average SRR difference, with headway difference (-.1249).
ADDITIONAL PROCEDURES RECOMMENDED: Rank correlations.</td></tr>
<tr><td>BRAKE APPLICATIONS</td><td>1. ANOVA OF RAW SCORES:
<table><tr><th></th><th></th><th colspan="4">MEANS</th><th></th></tr><tr><th>Test</th><th>N</th><th>8-Hr</th><th>4-Hr</th><th>S.T.</th><th>WEDMK</th><th>α</th></tr><tr><td>PRE</td><td>289</td><td>40.8</td><td>44.5</td><td>44.2</td><td>38.3</td><td>.0000</td></tr><tr><td>POST</td><td>292</td><td>38.0</td><td>41.0</td><td>42.7</td><td>38.6</td><td>.0121</td></tr><tr><td>F-UP</td><td>106</td><td>35.2</td><td>39.3</td><td>37.8</td><td>35.8</td><td>.1154</td></tr></table>2. ANOVA OF POST-PRE DIFFERENCES:
<table><tr><th>N:</th><th>8-Hour</th><th>4-Hour</th><th>S.T.</th><th>WEDMK</th><th>α</th></tr><tr><td>166*</td><td>-1.95</td><td>-4.09</td><td>-3.51</td><td>+.16</td><td>.1673</td></tr></table>*NOTE: This is considerably lower than pre- or post-test N's because BA's were not measured on all runs: however group N's are approximately equal.</td><td>PRODUCT-MOMENT CORRELATIONS WITH SBTW SUM SCORE (MAX=21) AND AVERAGE HEADWAY: Significant at α = <.05 :
Pre-test BA with pre-test SBTW sum (-.1301).
Post-test BA with post-test headway (-.1697).
ADDITIONAL PROCEDURES RECOMMENDED: Correlation with SBTW items from left turn/arterial sequences only.</td></tr>
<tr><td>INDEX NUMBERS</td><td>1. ANOVA OF THE ABSOLUTE VALUES OF THE STANDARDIZED LOG TRANSFORMATION OF AVERAGED INDEX NUMBERS:
<table><tr><th></th><th></th><th colspan="4">MEANS</th><th></th></tr><tr><th>Test</th><th>N</th><th>8-Hr</th><th>4-Hr</th><th>S.T.</th><th>WEDMK</th><th>α</th></tr><tr><td>PRE</td><td>278</td><td>.84</td><td>.76</td><td>.77</td><td>.82</td><td>.8590</td></tr><tr><td>POST</td><td>280</td><td>.85</td><td>.90</td><td>.65</td><td>.83</td><td>.0561</td></tr></table>2. NON PARAMETRIC TESTS OF THE DIFFERENCE BETWEEN PRE-TEST AND POST-TEST DISTRIBUTIONS OF AVERAGED INDEX NUMBERS:
<table><tr><th></th><th>N_1
N_2</th><th colspan="2">Significance Levels:</th></tr><tr><th></th><th></th><th>Mann-Whitney Rank</th><th>Kolmogoron-Smirnoff</th></tr><tr><td>Whole Sample</td><td>263
251</td><td>.5914</td><td>>.2</td></tr><tr><td>8-Hour</td><td>62
54</td><td>.4222</td><td>>.2</td></tr><tr><td>4-Hour</td><td>74
77</td><td>.4388</td><td><.25</td></tr><tr><td>S.T.</td><td>55
54</td><td>.2483</td><td>>.2</td></tr><tr><td>WEDMK</td><td>72
66</td><td>.7203</td><td>>.2</td></tr></table>ADDITIONAL PROCEDURES RECOMMENDED: Two-sample comparisons of individual index numbers.</td><td>PRODUCT-MOMENT CORRELATIONS OF THE ABSOLUTE VALUES OF THE STANDARDIZED LOG TRANSFORMATIONS OF EACH INDEX NUMBER, AT PRE-, POST- AND FOLLOW-UP TEST, WITH SBTW SUM SCORE (MAX=21), FREEWAY SUB-SCORE (MAX=9) AND AVERAGE HEADWAY: Significant at α = <.05:
Pre-test pre-stress I.N., with pre-test freeway sub-score (.1522).
Follow-up recovery I.N., with follow-up SBTW Sum (-.2633).</td></tr>
</table>

standardized scores (Z scores), which transform each observation in terms of its deviation from the mean. The absolute values of these were used to give a unidirectional comparison. Analysis of variance then clearly demonstrated that the extent to which the groups deviated from the overall mean was much more similar at post-test than at pre-test. Differences between groups in the amount of change in deviation from pre-test to post-test were then examined, and found to be non-significant. Only one correlation between the absolute value of Z scores and the subjective measures came close to significance; this was with average headway in the follow-up test. This suggests that those who keep more headway deviate less from the average time taken to complete the route.

Steering reversal rate (SRR) is a component of the index number, but it was also analysed separately. Between-group differences were not found at the 5% level or better, except on one pre-test sequence; raw scores and post-pre difference scores were both examined. Pre-test average SRR correlated positively with headway difference scores (post-pre), while post-pre differences in average SRR correlated negatively with headway difference scores. These correlations imply that drivers with higher initial SRR's improved their headway after taking the course more than those with low SRR's.

Brake Applications (BA) would tend to show a reduction if drivers were anticipating hazards and positioning their vehicles more advantageously. Significant differences were found between treatment groups on pre- and post-test, but significance was lower at post-test. Differences calculated as post-test score minus pre-test score suggest that the three treated groups improved, while the WEDMK group did not; however the significance of this result is only about 17%. Two correlations support subjective test findings, although the coefficients are low: those with fewer brake applications had better

SBTW sum scores on the pre-test, and allowed more headway on the post-test*.

The Index Numbers (IN) are composites of speed and steering reversals for the three freeway sequences of the test. The three numbers were averaged, and their distribution on the pre-test and post-test were compared using non-parametric two-sample techniques. Strongly significant differences were not found for the whole sample or individual treatment groups. Most of the subsequent analysis was directed towards deviation from whole sample or treatment group norms. Log transformations were performed to overcome skewness, and then transformed again into standardized scores (Z scores). The absolute values of the Z scores were used as in the Total Time analyses. On this basis, the treatment groups differed significantly at post-test, with the Self-Teach group showing the lowest deviation. Using the same transformations, two correlations were found to be significant at the 5% level: deviation increased with the score on SBTW freeway items at pre-test; and (more strongly) deviation decreased as SBTW sum score increased at follow-up. Extreme IN's may thus identify poor SBTW scorers.

These analyses are, to a great extent, exploratory. Moreover, no attempt was made in the design to relate objective measures to course content or subjective measures; hence the difficulty of direct comparisons. The data base contains much information which might be used to relate this type of measurement to the evaluation problem in more detail. At present, the OBTW data is marginally supportive of the SBTW results in some important areas. In addition, the data base provides a unique opportunity to pursue refinement of the OBTW measures for a variety of other purposes; heretofore, these measurements have not been applied under experimental constraints except on a much smaller scale.

*Brake applications are more pertinent to non-freeway driving, and it would be worthwhile to pursue more detailed comparisons with the subjective items from the left turn/arterial sequences.

Subjective Behind the Wheel Test

The SBTW test contains three replications of 9 items which may be used as a simple additive sum score, and three subjective estimates of car following time. Initial analyses were carried out for the sum score and for the average of the headway observations. In addition, an abbreviated sum score was derived by eliminating the three replications of two items which were often unscored because of light traffic conditions*. This sum score (maximum 21) correlated with the sum score for all items at .945 or better on the pre-test, post-test and follow-up test data. The abbreviated sum score was used in all subsequent analyses.

Table 5.12 shows the unadjusted means and ANOVA significances for the two types of data at pre-test, post-test and follow-up test, and for three difference scores. In contrast to the other tests, these show very marked improvements between pre-test and post-test. The sum score means are shown graphically in Figure 5.5, which also includes the sum score means for all items. It seems that the post-test pattern was retained at follow-up test, with an apparent further appreciable increase by the Four-Hour group. However, when the change from post-test to follow-up test is examined for those whose observers were the same on both tests, the three trained groups show small increases, and the WEDMK group an increase of 1.5. It will be noted that the differences between the groups in this small sample are significant only at the 23% level.

An even smaller sampling of those matched on observers at pre-test and follow-up test show a pattern of increase, in the same order as the intensity of training, which is significant at the 5% level.

The headway data also seem to reflect the different levels

*These items were: "Did driver make any attempt to improve space cushion?"; and, "Did driver keep his wheels straight ahead while waiting to turn left?".

TABLE 5.12
Means by Treatment Group for Key Subjective BTW Test Scores

Sum Score (Max=21)

TEST DATA:	TOTAL N	MEANS 8-HOUR	4-HOUR	SELF-TEACH	WEDMK	ANOVA α
Pre-Test*	428	9.64	9.98	9.35	10.0	.4643
Post-Test*	428	13.28	12.94	10.26	9.57	.0000
Follow-up*	81	13.24	14.06	10.05	10.79	.0002
Post - Pre** Difference	428	4.93	3.59	1.25	-.37	.0000
F-Up - Post** Difference	81	-.10	-.18	-.58	1.5	.2254
F-Up - Pre*** Difference	62	3.67	3.36	1.14	.88	.0449

Car following time (headway) in seconds averaged over 3 observations.

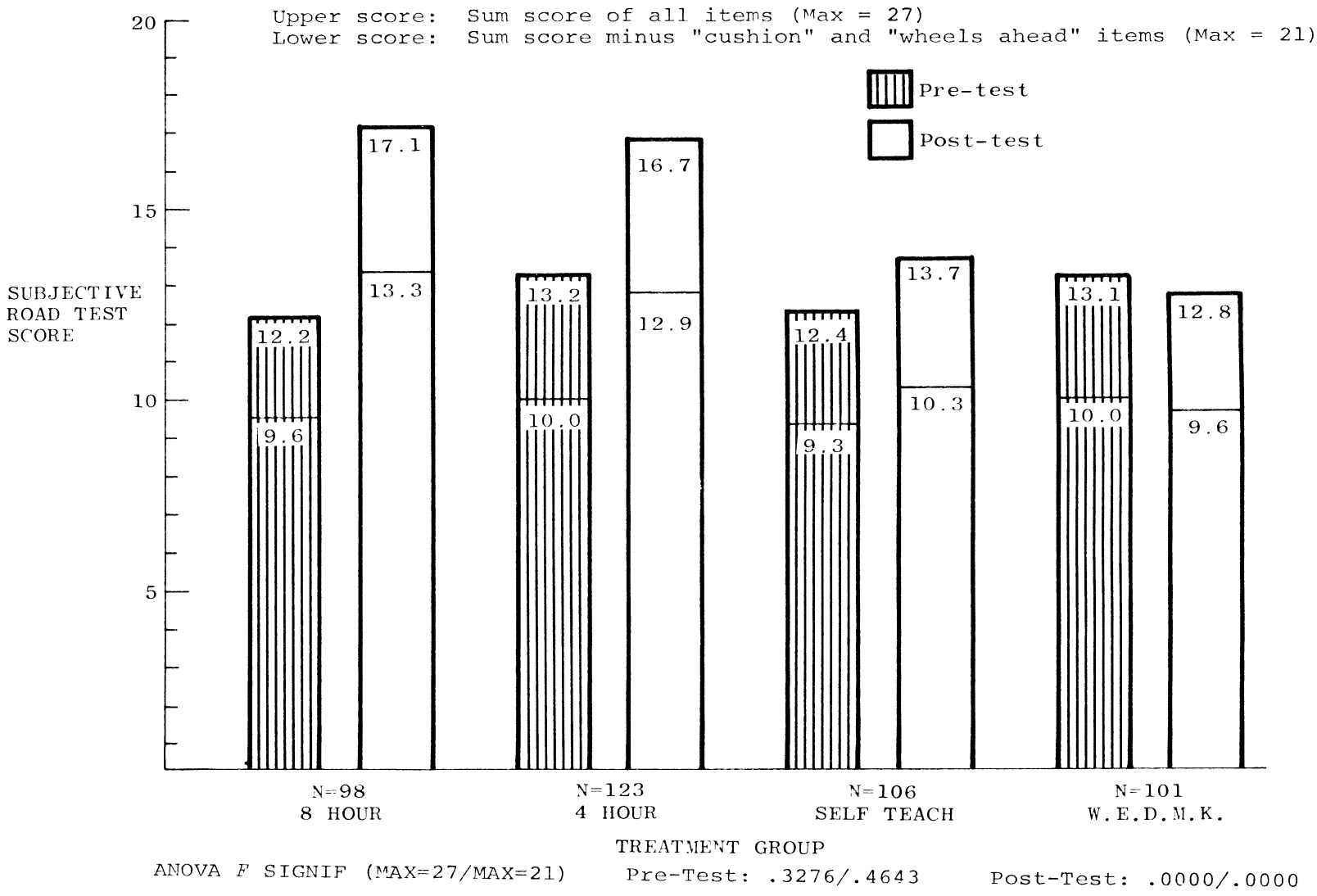
TEST DATA:	TOTAL N	MEANS 8-HOUR	4-HOUR	SELF-TEACH	WEDMK	ANOVA α
Pre-Test*	419	2.58	2.54	2.42	2.63	.0224
Post-Test*	419	2.82	2.68	2.62	2.56	.0022
Follow-up**	80	2.72	2.82	2.25	2.45	.0007
Post - Pre* Difference	419	.25	.24	.19	-.08	.0007
F-Up - Post** Difference	79	-.08	.21	-.24	-.03	.0969
F-Up - Pre*** Difference	58	.24	.54	.15	-.07	.1471

*Analysis confined to subjects having same observer pre- & post-tests.

**Analysis confined to subjects having same observer post- and follow-up tests.

***Analysis confined to subjects having same observer pre- and follow-up tests.

NOTE: Only subjects pre- and post-tested by the same observer have been included.



ANOVA *F* SIGNIF (MAX=27/MAX=21) Pre-Test: .3276/.4643 Post-Test: .0000/.0000

FIGURE 5.5
Subjective Road Test Mean Sum Scores, by Treatment Group, For Pre-Test and Post-Test

of training intensity. Post-pre differences are shown in Figure 5.6, which gives pairwise *t* significance levels for each treatment group. Table 5.12 shows that this pattern became erratic between post-test and follow-up test. Again, the three treated groups increased noticeably between pre-test and follow-up test, especially the Four-Hour group, but the significance level between groups is only 15%.

These advantageous initial results on SBTW Sum Score and car following time clearly warranted more detailed analysis. In particular, the use of difference scores tends to compound any unreliability in the raw scores. In order to examine the net effect of training more accurately, one-way analyses of covariance were used. For both types of data, models were built using post-test score as the dependent variable with pre-test score as the covariate. Follow-up test score was also used as a dependent variable in two models for which post-test and pre-test scores were the covariates. In this setting analysis of covariance procedures essentially consist of the construction of a regression equation to predict post-test scores from pre-test scores, together with the examination of residual variance not explained by the equation. Treatment effects, or other influences on treatment group mean scores, would not be explained by the equation, and if large enough, could be reflected as significant differences between group means after adjustment for the covariate.

The results of the six models are summarized in Table 5.13 This table gives three *F* test significance levels. The first is for the difference between adjusted means; the second covers an assumption of analysis of covariance, namely that the slopes of the regression lines for each of the treatment groups do not significantly deviate from parallel; and the third tests the hypothesis that the slope of the covariate is not significantly different from zero. In this situation, slopes which differ at the <.05 level would suggest that the model is inappropriately used; none of the models have levels lower than .08. A

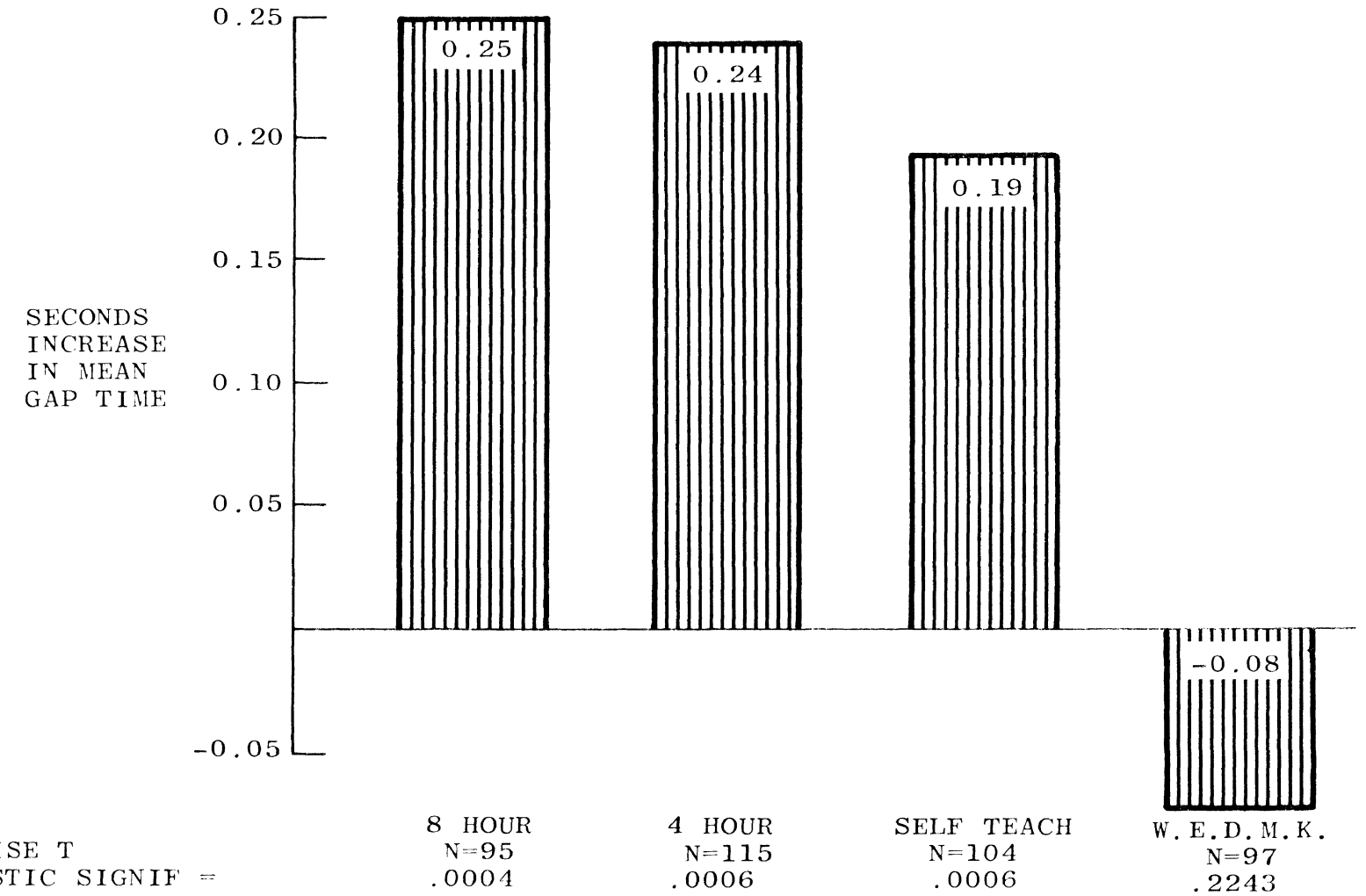


FIGURE 5.6
Post-Test Minus Pre-Test Differences in Car Following Time, by Treatment Group

TABLE 5.13
Summary of One-Way Covariance Models for SBTW Sum Score (Max=21) and Average Headway (Car Following Time

VARIABLE	COVARIATE	N	ADJUSTED MEANS 8-HR	4-HR	S.T.	WEDMK	*F* SIGNIF ADJUSTED MEANS	*F* SIGNIF EQUALITY OF SLOPES	*F* SIGNIF COVARIATE	PAIRWISE COMPARISONS: *t* SIGNIFICANCE 8-HR vs 4-HR	8-HR vs S.T.	8-HR vs WEDMK	4-HR vs S.T.	4-HR vs WEDMK	S.T. vs WEDMK
SBTW Sum, Post-test	SBTW Sum, Pre-test	*428	13.33	12.86	10.44	9.44	.0000	.0840	.0000	.3406	.0000	.0000	.0000	.0000	.0353
SBTW Sum, Follow-up	SBTW Sum, Post-test	**81	12.60	13.08	10.46	11.76	.0515	.2177	.0001	.6085	.0227	.3333	.0101	.1746	.1543
SBTW Sum, Follow-up	SBTW Sum, Pre-test	***62	13.77	14.07	10.26	10.91	.0003	.9663	.0001	.7553	.0008	.0035	.0007	.0026	.5220
Seconds of Head-way, Post test	Headway, Pre-test	*411	2.81	2.70	2.64	2.53	.0007	.1053	.0000	.1075	.0128	.0001	.3301	.0094	.1062
Seconds of Head-Way, Follow-up	Headway, Post-test	**79	2.65	2.82	2.27	2.49	.0026	.2263	.0060	.2581	.0075	.2032	.0003	.0185	.1019
Seconds of Head-way, Follow-up	Headway, Pre-test	***58	2.73	2.80	2.28	2.45	.0108	.3302	.5194	.6913	.0095	.0765	.0046	.0458	.3268

*Analysis confined to subjects having same observer pre- and post-tests

**Analysis confined to subjects having same observer post- and follow-up tests.

***Analysis confined to subjects having same observer pre- and follow-up tests.

covariate significance of >.05 would indicate that the model is unnecessary (but not harmful), and that the means are essentially unaffected; the sixth model is in this position. The table also gives pairwise *t* test significances for all possible pairs of treatment groups.

For the SBTW sum score, these analyses strongly confirm the ANOVA results for the basic post-test/pre-test comparison. The adjusted means reflect the intensity of treatment, and between-group significance is very high. All pairs of the treatment groups, except Eight-Hour versus Four-Hour, are significantly different. In addition, the pattern of follow-up test scores remained essentially the same. Some dominance of the Four-Hour group at follow-up testing, which was ambiguous in the ANOVA analysis, is now evident in relation both to post-test and to pre-test scores, although significance is marginal for the former. However, on none of the models does the Four-Hour group differ significantly from the Eight-Hour group. The WEDMK group is somewhat higher than the Self-Teach group at follow-up, but pairwise comparisons do not show significant differences between these two groups. The remaining pairwise comparisons at follow-up appear to support the retention of a training effect by the Eight-Hour and Four-Hour groups, but not completely; the WEDMK group mean is high enough in the follow-up/post-test model not to differ significantly from either of those two groups.

The headway covariance models reveal a very similar pattern to the sum score models: a highly significant difference exists between the post-pre adjusted means, with the order and pairwise comparisons supporting the apparent effectiveness of the Eight-Hour and Four-Hour courses; significant differences in the follow-up test adjusted means again favor the Four-Hour group to an extent not supported by differences between it and the Eight-Hour group; pairwise comparisons at follow-up suggest that the formally trained groups are distinct from the Self-Teach and WEDMK groups; and the WEDMK group is again higher than the Self-Teach group, but not to a statistically significant extent.

Altogether, compared to the ANOVA analyses, the covariance models are less ambiguous, and more soundly supportive of the effects of the Eight-Hour and Four-Hour courses. However, it is important to examine the possibility that these results are explained by something other than intensity of training. The experimental design was devised and validated (see Section 4.2) to ensure a minimum of external bias in the program results. However, subjective tests are liable to internal bias attributable to the observers. Such problems may arise first in the form of inconsistencies between observers in their manner of scoring subjects in general, or certain groups in particular.

The consequences of simple inconsistencies in observational style were minimized by the restriction of analyses to cases with matched observers, by the random allocation of observers to treatment groups, and to some extent by the use of difference scores and covariance analysis. In fact, substantial differences were found between observers in the mean of all the scores they awarded. More problematical were significant differences in the average difference between a given observer's pre-test and post-test scores. It was yet more important to examine the possibility of interaction between treatment groups and individual observers; in other words, to examine inconsistencies in the average amount of improvement or deterioration recorded by the observers for each treatment group*. At an exploratory level, mean scores were examined for all combinations of observer and treatment group. Some irregularities were noted, especially for one observer who characteristically awarded much higher increases to the Eight-Hour and Four-Hour groups, compared to the other observers. Therefore, it was decided to build regression models for SBTW sum score, and average headway, with treatment group and observer as predictors to be considered independently, and conditional upon each other. By the use of dummy variable techniques, it is possible to obtain significance

*It will be recalled that it was not possible in practice to conceal treatment group membership from the observers.

tests for categorical factors such as these, to estimate the percentage of variance explained by each, and to obtain some evidence of interactions.

Table 5.14 gives the results of the dummy variable regression analysis for SBTW sum score. Treatment and observer factors were significant both independently and conditionally. However, as independent factors, treatment explained approximately 14% of the variance, whereas observer explained only 4%.

TABLE 5.14
Conditional Variance Table for Dummy Variable Regression of Treatment and Observer Effects on SBTW Sum Score (Max = 21)

Dependent variable: post-pre difference score N=428*

MODEL NUMBER	R^2	SOURCE	D.F.	SUM OF SQUARES	MEAN SQUARES	*F* STAT	SIGNIF α=
1	.139	Treatment, alone	3	1123.2	374.4	22.69	.0000
3-2	-	Treatment, conditional on observer	3	1079.9	360.0	22.36	.0000
2	.042	Observer, alone	7	344.6	49.3	2.67	.02
3-1	-	Observer, conditional on treatmnt	7	301.3	43.0	2.67	.02
1		Error	424	6985.0	16.5	-	-
2		Error	420	7763.6	18.5	-	-
3	.176	Error	417	6683.8	16.1	-	-
ALL	-	Total	427	8108.2	-	-	-

*Only cases with the same observer on pre- and post-test were used.

Moreover, the R^2 with both factors in the equation (model #3 in the table) was 17.6%, barely less than the combined R^2's of models #1 and #2; this, and the closeness of the *F* statistics suggests very little interaction. This regression analysis also raises the question of the meaningfulness of highly significant differences in treatment effects which explain only 14% of the variance. It will be recalled that an earlier study of the SBTW scores showed that a regression model containing age, salary grade and driving experience as predictors, could only explain about 7% of the variance. This result is thus quite meaningful relative to factors which are normally linked to driving style and performance, as well as to observer inconsistencies.

An identical procedure was followed with average headway, and the results of this are given in Table 5.15. In this case observer effects were not significant, independently or conditionally. However, treatment was again significant on both counts and once more, interaction appears to be negligible. As an independent factor, treatment explained 4.2% of the variance in average headway. This suggests much caution in interpreting the greater increase in headway achieved by the Eight-Hour and Four-Hour groups.

The other form of internal bias which may have influenced the subjective scores is much more elusive. It is possible that as a group the observers may consciously or unconsciously have allowed treatment group membership to influence their scoring régime. By and large, the observers believed themselves to be impartial. Unfortunately, it is almost impossible to detect bias which was systematically partisan with respect to treatment groups, and which was consistent among the observers. The only basis for comparison which exists is open to serious sampling objections: it is that Eight-Hour participants who completed only one or two of their four sessions have had the same gross amount of training as the Four-Hour participants who completed one or both of their allotted sessions. There were 12 from the Eight-Hour group who took only two sessions,

TABLE 5.15
Conditional Variance Table for Dummy Variable Regression of Treatment and Observer Effects on Car Following Time (Headway)

Dependent variable: post-pre difference score N=411*

MODEL NUMBER	R^2	SOURCE	D.F.	SUM OF SQUARES	MEAN SQUARES	*F* STAT	SIGNIF α=
1	.041	Treatment, alone	3	266.38	88.79	5.74	.0007
3-2	-	Treatment, conditional on observer	3	245.88	81.96	5.30	.002
2	.02	Observer, alone	7	136.93	19.56	1.23	.2868
3-1	-	Observer, conditional on treatmnt	7	116.43	16.63	1.08	>.25
1		Error	407	6298.6	15.48	-	-
2		Error	403	6428.0	15.95	-	-
3	.058	Error	400	6182.1	15.46	-	-
ALL		Total	410	6565	-	-	-

*Only cases with the same observer on pre- and post-test were used.

and one participant who took just one. An analysis of variance suggested that there was no significant difference in SBTW sum score between these 14 Eight-Hour participants, the 148 who completed the Four-Hour course, and the 23 who took only one of the Four-Hour sessions. The mean score for the Eight-Hour participants who took only two sessions, was slightly higher than that for Four-Hour participants who took the same amount of training. Strong bias in favor of the Eight-Hour course could be expected to show up in this comparison, but it is also true

that a similar bias towards both the Four-Hour and the Eight-Hour courses would not. From the analyses in general, and from an extensive debriefing of the observers immediately after the end of the program, it is reasonable to assume some partiality towards the more intensively trained groups. However, the increases in subjective scores between pre- and post-test are almost certainly too large to be explained by this alone.

The implication of measured improvements in SBTW sum scores is that the more intensive treatments are capable of bringing about increased performance of certain behaviors which are logically related to accident-free driving; moreover, recipients of those treatments are, on the average, able to demonstrate a similar performance some months after the end of the program. This conclusion may be extended to improvements in average headway only on much more marginal statistical grounds; however, if this finding can be substantiated in later progroms, it will represent a most significant contribution to the upgrading of of unexceptional adult drivers.

5.3 PROSPECTS FOR ULTIMATE EVALUATION

On the basis of recorded frequencies of collisions and the sample sizes finally achieved, rather lengthy periods are required to detect small percentage changes with statistical significance. Specifically, by combining the treatment groups into two categories (Eight-Hour plus Four-Hour, and Self-Teach plus WEDMK), comparisons can be made between approximately 280 participants in each. The SOS data records a rate for all collisions which averages .07 per driver year*. At this rate, and at a confidence level of 5%, there is a 90% probability of detecting a 50% reduction in collisions in two years, or a 45% reduction in four. There is a 75% and 50% chance of detecting, respectively, 40% and 30% reductions in two years, or 35% and 25% reductions in four.

*In Michigan, all male drivers between the age of 21 and 65 years average .08 recorded collisions per year, according to state files.

The TARIC data, being restricted to employes above Grade 9, must use smaller samples. Again by combining the groups, comparisons may be made between samples of approximately 150 formally trained, and 150 self-taught or untrained participants. A dollar loss rate reduction of 25%, for example, based on .3 claims per driver year, could be detected with 85-90% probability in two years, also with a confidence level of 5%. In follow-up periods, careful consideration must be given to any changes in claims procedures which may have occurred.

If follow-up study periods are insufficient for the size of changes occurring in the "ultimate" data, consideration will be given to alternatives to conventional comparisons of mean rates. One possibility is to compare post-program accident free periods using non-parametric statistical techniques.

6.0 CONCLUSION AND RECOMMENDATIONS

This evaluation has examined the effectiveness of a driver improvement project, known as the Ford Employe Skilled Driving Program, which was designed to yield information about the amount of retraining appropriate for populations of unexceptional adult drivers. In our opinion, this program has demonstrated that it is feasible to operate fairly large driver programs in a company setting under experimental controls essential for evaluation, even though the present program experienced considerable difficulties. Our conclusions about the major aspects of the program are as follows:

Sampling

Approximately half (1052) of the 20% random sample of all Detroit area salaried employes agreed to participate. Attrition from this group was substantial, with 750 employes attending at least a pre-test, and only 560 participating through the post-test. Higher participation levels are much to be desired, but in this case it could not be shown from an analysis of biographical and driving record characteristics that the volunteers who took part were significantly different, as a group, from those who cancelled or dropped out. More importantly, the sample trained could not, for practical reasons, include any non-volunteers; this tends to limit the generalizability of the status and years of service with company, suggested that those who took part did not markedly differ, as a group, from Ford salaried employes as a whole. However, it is not known the extent to which the participant sample, with its tendency toward the middle salary grade and age ranges, can represent the general driving public.

Training Procedures

These were devised using what Ford Motor Company and its consultants considered to be the most appropriate content, methods and materials available for unexceptional drivers of

widely differing ages and backgrounds. Six heterogeneous instructors were selected to encourage a range of teaching styles. Comparisons were made possible by the experimental design between four treatment groups, corresponding to three intensities of training and control. Two of these groups underwent classroom and in-car instruction, and were thus vulnerable, from an evaluation standpoint, to inconsistencies in the allocation of instructors. Moreover, three of the original six instructors resigned during the course of the study and two replacements were hired and given on-the-job training. Analyses of class records showed that instructors were close to randomly distributed over the in-car sessions, but not the classroom sessions. While the intention was not to evaluate in detail the training situations, it should be noted that there was a great deal of variability over time in the quality of the training procedures.

Testing Procedures

A set of measures was developed from the evaluation model, and the most appropriate available instruments were employed. Much was learned in the program about the application of tests to this kind of evaluation. Cognitive tests are probably not a problem unless program content becomes much more technical in nature. The classroom tests in the perception and decision-making areas suffered a variety of problems, including some mechanical difficulties. Better measures in these areas are essential if firm benefits are to be obtained from causal chain methodology. Their inclusion was worthwhile, however, for the information they yield on the response of different subgroups of adults to the training and testing situations.

The objective in-car measures had not heretofore been applied on a large scale under experimental constraints; such applications are essential for their development. The present data are difficult to interpret in the context of the evaluation design, but as intended, considerable potential exists for a posteriori analyses. The subjective measures were appropriate for this program, but they need to be made less ambiguous for inexperienced observers. Some of them, notably headway, could

perhaps be augmented with parallel objective devices, even if they were used only to improve the training of observers. The use of instructors as observers within the same program is definitely to be discouraged.

Results

The most favorable results were provided by the subjective rating of behind-the-wheel performance. Within this test, the mean sum scores for the performance of behaviors logically related to accident-free driving increased between pre- and post-test increased 40%, 30%, 10% respectively, for the Eight-Hour, Four-Hour, and Self-Teach groups, compared to slight decrease for the control group. The three treated groups appeared to increase car following time (headway) by between one-fifth and one-quarter of a second. The sum scores were substantiated using several statistical techniques; the improvements in car following time (headway) were marginally supported by similar analyses. However, it is probable that part, but not all, of the differences recorded are explainable by observer bias, as it proved impossible to conceal the treatment group of subjects from the observers. The objective measures offered only a few weak indications of treatment group differences.

Of the remaining tests, the Unusual Uses technique showed small improvements in the ability of the Eight-Hour and Four-Hour participants to think imaginatively about driving, but the Perception of Hazards tests was unable to detect group differences in performance resulting from the course. The smallness of differences in both the Content Acquisition and Driving Knowledge portions of the Written Test suggests that the transfer of key information from programs such as this may well be achieved with little instructional effort. However, all groups showed unexpected deficienceis in general driving knowledge, even though all participants were given a copy of the new Michigan driver's manual.

The survey data reveals generally high approval of the concept of advanced driving training for adults and a range of

preferences as to how and when the respondents would prefer to undergo it. There was considerable endorsement of behind-the-wheel training and of audio-visual instructional materials, but the overall demand was for some extensions in content and for a variety of learning situations.

Although the test results improved with the intensity of training, the greatest increase occurred between the Self-Teach and Four-Hour groups, that is between those who were not formally trained and those who were. Support exists in the data for conducting several hours of training, but it can not be shown that eight hours were substantially better than four.

The HSRI data base on the Employe Skilled Driving Program is a valuable resource for further study of adults' responses to such offerings. It is a product of procedures more rigorous than are generally applied to programs involving large numbers of "average" adult drivers.

HSRI's key recommendation for future activity in the area of adult driver re-training is that, given a reasonable definition of appropriate instructional content, much attention should be directed to identifying "target groups" of drivers for a variety of instructional procedures. For large populations of unexceptional drivers, training programs must be limited to a few hours; under these circumstances, and given the state-of-the-art of testing, it is unrealistic to expect to develop truly individualized instruction for all drivers in the near future. However, with due attention to the acceptability of various training approaches, including those used in this program, and to the compliance to driving standards which is implicit in different training situations, the characteristics of such traget groups should begin to emerge.

Because of the great number of factors which influence driver performance, rather small amounts of change resulting from improvement and re-training programs are meaningful if they can be supported by a rigorously controlled evaluation. Indeed, dramatic amounts of change have not been proven in the

literature for large numbers of drivers. We would contend that an evaluation of this complexity is essential to provide accurate information on the amounts and types of training which can be supported as efficient on a large scale. The present study will remain active for several years, as accident and loss data are accumulated and analyzed for the subjects.

This study completes Phase I of a program to disseminate retraining methodologies for unexceptional drivers, with special reference to applications in the driver licensing process. We recommend a continuing effort to assess the effectiveness of these methodologies as they are developed, and before they are implemented.

REFERENCES

[1]B. J. Campbell, Commentary on the Removal From the Driving Population of Drivers with Recent Accident and Violation Records. Signal 99, North Carolina Governor's Highway Safety Program, Volume 1, Number 2, Raleigh, Spring 1972.

[2]Albert Burg, "The Stability of Driving Record Over Time", Accident Analysis and Prevention, Volume 2, May 1970, Pergamon Press, London, 1972. pp. 57-65.

[3]J. A. Whittenburg, R. F. Pain, Robin McBride, Joseph Amidei, Driver Improvement Training and Evaluation: Initial Development. Final Report. The American University, Development and Training Research Institute, Washington, D.C., May 1972.

[4]Martin E. Lee, "Detecting the Effectiveness of Driver Retraining Among Detroit Car Commuters", Universities Transport Study Group Conference. Proceedings. University College, London. January 3-4, 1973. January 1973.

[5]The agencies contracted (and principal investigators) were:

New York University (L. Brody)
American University (W. A. Lybrand)
Dunlap and Associated (G. E. Teal)
Institute for Education Development (J. L. Kennedy and R. L. Chapman)

[6]H. H. Harman, D. W. Seibel, M. Rosenfeld, and B. Shimberg, Evaluation of Driver Education and Training Programs, Highway Research Board, March 1969.

[7]J. O'Day, J. S. Creswell, Jr., J. A. Green, M. E. Lee, J. C. Marsh, IV, and S. Schultz II. Planning for the Concentrated Implementation of Highway Safety Countermeasures. Highway Safety Research Institute, University of Michigan, August 1971. 4 Volumes.

[8]K. McPherson and Cooper, Perception of Hazards Test. Illinois State University, Normal Illinois. Circa 1966.

[9]F. N. Platt, P. Gram and G. F. Hobday, "An Historical Review of the Highway Systems Research Car Instrumentation", Instrument Society of America Annual Conference, Houston, October 1969. Pittsburg, Pennsylvania. 1969.

[10]F. N. Platt, "Objective Measurements of Individual Driver Behavior", Automotive Engineers Congress, Detroit, Michigan, January 1964. Society of Automotive Engineers, New York, 1964.

[11]H. Scheffe, The Analysis of Variance, John Wiley, New York, 1959.

APPENDIX 1

COURSE OUTLINES

OUTLINE OF TRAINING SESSIONS FOR EIGHT AND FOUR HOUR COURSES

The eight hour course consisted of all four sessions; the four hour course consisted of the first and second sessions. Each classroom period was immediately followed by a one hour Behind-The-Wheel training period. A car and instructor were assigned to each two participants, providing 30 minutes of driving and 30 minutes of observation for each participant.

Outline of Classroom Sessions (from Instructor's Guide)

First Classroom Period

A. Subject - (First Classroom Period)

1. Course Orientation----------15 minutes
2. Showing of "9 out of 10" Filmstrip----------10 minutes
3. Discussion of Filmstrip---------- 5 minutes
4. Introduction of "Decision Pattern" Filmstrip-------- 5 minutes
5. Showing of Filmstrip----------10 minutes
6. Discussion and Response----------15 minutes

B. Objectives

1. The driver will be able to identify and describe the characteristics of drivers who think they are above average drivers.
2. The driver will be able to make a self-appraisal of his own driving.
3. The driver will be able to verbalize and explain fully the six major steps of "The Decision Pattern."

C. Materials and Equipment

1. Chalkboard (chalk and eraser)
2. Sound Slide Film Projector and Screen
3. Extension Cords and Convenient Outlets
4. Classroom
   a. Well lighted and ventilated.
   b. Equipped with participant writing facilities--paper and pencils.

D. Teaching Aids

1. Filmstrip - "9 out of 10" - with accompanying guide - Ford Motor Company.
2. Filmstrip - "Decision Pattern" - with accompanying Instructor's Guide - Ford Motor Company.

Second Classroom Period

A. Subject - (Second Classroom Period)

1. First Hour Classroom Review---------- 5 minutes
2. Introduction of 16 mm Film on Rural Driving--------- 5 minutes
3. Showing of Film----------10 minutes
4. Discussion of Film----------10 minutes
5. Introduction of Film on Commentary Driving---------- 5 minutes
6. Showing of Film----------10 minutes
7. Discussion of Film---------- 5 minutes
8. Course Summary and Driving Instructions----------10 minutes

B. Objectives

1. The driver will be able to make accurate decisions when considering relationships with environmental elements of the roadway and with other vehicles.
2. The driver will be able to execute maneuvers involving speed decisions more smoothly, accurately and efficiently.

C. Materials and Equipment

1. Chalkboard (chalk and eraser)
2. 16mm Motion Picture Projector
3. Extension Cords and Convenient Outlets
4. Classroom
   a. Well lighted and ventilated.

b. Equipped with participant writing facilities--paper and pencils.

D. Teaching Aids

1. Motion Picture - "Rural Driving" - Ford Motor Company.

2. Motion Picture - "Commentary Driving" - Ford Motor Company.

Third Classroom Period

A. Subject - (Third Classroom Period)

1. Introduction to 16 mm Film, Driving in Traffic------ 5 minutes

2. Showing of Film----------10 minutes

3. Discussion of Film----------15 minutes

4. Introduction to Filmstrip, Strategic Positioning---- 5 minutes

5. Showing of Filmstrip----------10 minutes

6. Discussion and Response----------10 minutes

7. Class Review and Driving Instruction---------- 5 minutes

B. Objectives

1. The driver will be able to make accurate decisions involving proper positioning relationships with other vehicles and with pedestrians.

2. The driver will be able to execute maneuvers involving space and positioning relationships smoothly, accurately and efficiently.

C. Materials and Equipment

1. Classroom

a. Well lighted and ventilated.

b. Equipped with participant writing facilities--paper and pencils.

c. Equipped with chalkboard (chalk and eraser).

d. Extension cords and convenient outlets.

D. Teaching Aids

1. Motion picture - "Driving in Traffic" - Ford Motor Company.

2. Filmstrip - "Strategic Positioning" - Ford Motor Company.

Fourth Classroom Period

A. Subject - (Fourth Classroom Period)

1. Introduction to 16 mm Film, Critical Driving Patterns---------- 5 minutes

2. Showing of Film----------10 minutes

3. Discussion of Film----------15 minutes

4. Introduction to Filmstrip, Emergency Problems------- 5 minutes

5. Test on Filmstrip (Response)----------10 minutes

6. Class Summary and Driving Instructions----------15 minutes

B. Objectives

1. The drivers will be able to develop habit patterns that will enable them to cope with the loss of one or more of the vital perceptual-vehicle control relationships occurring in critical or emergency situations.

2. The drivers will be able to regain one or more of the control abilities decreased or lost because of environmental, vehicular or psychophysical factors.

C. Materials and Equipment

1. Classroom

a. Well lighted and ventilated.

b. Equipped with participant writing facilities--paper and pencils.

c. Equipped with chalkboard (chalk and eraser).

d. Extension cords and convenient outlets.

2. 16mm motion picture projector and screen.

3. 35mm sound slide film projector.

D. Teaching Aids

1. Motion picture - "Critical Driving Patterns" - Ford Motor Company.

2. Filmstrip - "Emergency Problems" - Ford Motor Company.

WRITTEN INSTRUCTIONS, OUTLINE, AND SAMPLE TEST SHEET FOR SELF-TEACH COURSE

FORD EMPLOYE SKILLED DRIVING PROGRAM

SELF-TEACH COURSE

The Self-Teach Course is a combined self-appraisal and self-instructional course designed to help you upgrade your driving skills. Don't try to go through all the material at one time. Rather, pace yourself. Absorb each lesson and then begin applying what you've learned behind-the-wheel. Follow the enclosed outline sequence of materials. Spread your reading and practice driving time over no more than four weeks.

Learn the Decision Pattern and how to apply it; it is the basis of the Course and will aid you toward becoming an expert driver.

Volunteers in the Ford Skilled Driving Program also have available to them the use of a modern Audio-Visual Library. All of the materials in this packet have been programmed for use in the library. Program participants can work in privacy and at their own pace, using the latest equipment and visual aids.

For further information, please telephone 32-27047.

Phil Gram
Employe Skilled Driving Program

SELF-TEACH COURSE

STUDY SEQUENCE

1. "9 Out of 10" booklet and self-appraisal quiz.

2. The Decision Pattern

   - Workbook and multiple choice test.
   - Decision Pattern decal

3. Strategic Positioning

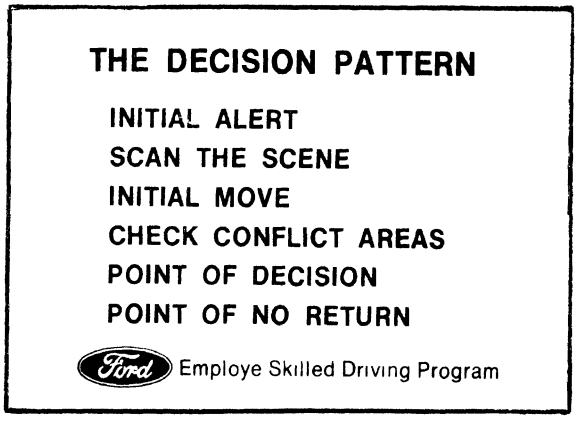
   - Workbook and multiple choice test.

4. Emergency Problems

   - Workbook and multiple choice test.

5. Michigan's driver manual, "What Every Driver Must Know," and multiple choice test. (80 questions).

Decision Pattern Decal



NOTES: The booklet and workbooks mentioned above are materials from Ford's American Road Skilled Driving Programs series (published by Helm Inc.).

Answer codes for all multiple choice tests were supplied with the instructions.

SAMPLE MULTIPLE CHOICE TEST SHEET FOR SELF-TEACH GROUP

CODE W-1/AS

1U. When you identify the location of the side street into which you have decided to make a left turn, you have received your:
a. initial alert
b. primary response
c. first positive warning
d. most important reaction

1L. Be aware of your initial alert as soon as possible. To get ready for the turn properly you need at least:
a. two blocks before turning
b. 75 feet before the intersection
c. one block before your turn
d. 50 feet from your planned turn

	A	B	C	D
1U				
1L				

2U. Scan the whole scene ahead following a left to right pattern in order to be able to locate and concentrate on:
a. your speed of approach
b. the important details
c. the location of crosswalks
d. all the traffic maneuvers

2L. Evaluate the important elements in the scene now, giving special attention to those elements which are:
a. possible hazards or conflicts to you
b. hidden by visual obstructions ahead
c. a part of the general layout
d. not included in your periodic checks

2U				
2L				

3U. Begin your initial move by first communicating your intention to turn to other drivers by:
a. sounding your horn & giving hand signal
b. flashing your brake lights rapidly
c. positioning your car on the center lines
d. signaling with your turn signal flasher

3L. Next, make your initial move by steering your car so that you position it properly. The best position is:
a. 1½ to 2 ft. right of the center line
b. 1/2 foot to the right of the center line
c. directly on the center line
d. about 1 foot left of the center line

3U				
3L				

4U. Now you must check the conflict areas ahead. To make this re-evaluation quick and accurate you should concentrate on:
a. not steering too close to center line
b. checking down the cross-street to the left
c. the conflict areas involved in your turn
d. traffic located behind visual obstructions

4L. At your "point of decision" you must decide to either continue your turn without stopping or to.
a. accelerate and abandon your turn
b. pull to the right & wait for an opening
c. stop and yield to oncoming traffic
d. stop before entering the intersection

4U				
4L				

5U. You must continue to scan the scene to the left, ahead, to the right and behind for new conflicts or possible hazards to your planned left turn throughout your approach to the intersection.

a. True b. False

5L. In addition to scanning the scene at the intersection ahead, you must also make periodic checks of:
a. the rate at which you are slowing down
b. your rear view mirror for hazards behind
c. your distance from the intersection
d. the speed of the oncoming traffic

5U				
5L				

6U. You should keep checking conflict areas both ahead and behind. If necessary, warn a "tailgating" driver by:
a. increasing your speed suddenly
b. waving your hand out the side window
c. sounding a long, steady horn signal
d. flashing your brake lights several times

6L. In this situation you have decided to stop and yield, you should pull ahead slowly and stop at a point:
a. about one car length from the crosswalk
b. just before entering the crosswalk
c. just past center of the intersection
d. just before center of the intersection

6U				
6L				

7U. Since you must stop in the intersection and yield to oncoming traffic, you should:

a. pull slightly over the center line
b. signal traffic behind you to pass
c. keep your wheels pointed straight ahead
d. turn your wheels to the right for safety

7L. While you are waiting to turn, re-check your rear view mirror again and check ahead for:
a. other vehicles passing in your lane
b. oncoming traffic hidden by a turning car
c. visual obstructions in the right lane
d. pedestrian activity in the next block

7U				
7L				

8U. As soon as the turning car has cleared you must re-check the major conflict area, which in this case is:
a. the side street you will enter
b. the area directly behind your car
c. the side street to the right
d. the lane of traffic directly ahead

8L. As you yield, check to make sure the side street is still clear of pedestrians and obstacles, and then begin to:
a. watch for a safe opening in traffic
b. move forward very slowly
c. turn your front wheels to the left
d. cross over the center line

8U				
8L				

9U. When you have a safe opening in traffic and begin your turn, you have reached your "point of no return". This means that you must continue your turn, as you are now committed to completing your maneuver.

c. True d. False

9L. Begin your turn and accelerate gradually. As you turn, be careful not to:

a. look down the street you are entering
b. cut the corner or loop out to the right
c. clear the intersection quickly
d. spend more time checking for pedestrians

9U				
9L				

10U. The main reason you must check the parked cars closely at this point is to be certain that no one is pulling out from the curb into your path.

c. False d. True

10L. Plan your turn and steer your car so that you will enter the new street:

a. just to the right of the center line
b. close to the outside curb
c. with your wheels pointed straight ahead
d. at the posted speed limit

10U				
10L				

ADVANCED DRIVER SERIES #1 - THE DECISION PATTERN

NAME _____

DATE _____ CLASS _____

AP-3

"WEDMK" GROUP (CONTROL GROUP): WRITTEN INSTRUCTIONS, AND SAMPLE TEST SHEET, ACCOMPANYING THE "WHAT EVERY DRIVER MUST KNOW" BOOKLET

FORD
EMPLOYE SKILLED DRIVING PROGRAM

WHAT EVERY DRIVER MUST KNOW

Michigan's official driver manual, "What Every Driver Must Know," provides pertinent information relative to laws and safety rules of the road. If you drive in Michigan, this manual is for you.

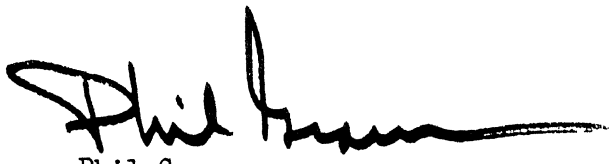
Read it and use it as reference. Share it with family and friends.

The manual has been programmed for self-instruction with a set of 80 multiple-choice questions. Take your time reading and digesting the information, then test yourself. (Answer code is provided below).

Employe Skilled Driving Program participants also have available the use of a modern Audio-Visual Library. All course materials have been programmed for use in the Library.

You can work in privacy and at your own pace, using the latest equipment and visual aids. There are more than 26 subjects from which to select.

For further information, please telephone 32-27047.



Phil Gram
Employe Skilled Driving Program

ANSWER CODE - "WHAT EVERY DRIVER MUST KNOW"

X - 1M Upper		X - 1M Lower		X - 2M Upper		X - 2M Lower	
1 - D	6 - A	1 - B	6 - A	1 - C	6 - D	1 - D	6 - C
2 - C	7 - B	2 - D	7 - C	2 - B	7 - A	2 - B	7 - A
3 - A	8 - D	3 - D	8 - D	3 - B	8 - B	3 - C	8 - A
4 - B	9 - B	4 - B	9 - C	4 - C	9 - D	4 - A	9 - D
5 - D	10 - A	5 - C	10 - D	5 - A	10 - A	5 - C	10 - B

X - 3M Upper		X - 3M - Lower		X - 4M Upper		X - 4M Lower	
1 - B	6 - A	1 - C	6 - B	1 - D	6 - C	1 - B	6 - A
2 - D	7 - C	2 - A	7 - D	2 - B	7 - A	2 - B	7 - B
3 - D	8 - D	3 - B	8 - B	3 - C	8 - A	3 - C	8 - D
4 - B	9 - C	4 - D	9 - A	4 - A	9 - D	4 - A	9 - A
5 - C	10 - D	5 - A	10 - C	5 - C	10 - B	5 - D	10 - A

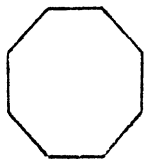
SAMPLE MULTIPLE CHOICE TEST SHEET FOR "WEDMK" GROUP

Michigan Driver's License Manual
Pages 18-27 (UPPER); 27-30 (LOWER)
Programed by INSTRUCTIVE DEVICES, INC.
Pawtucket, Rhode Island

NAME _____ X-1/M UPPER
CLASS _____
DATE _____ SCORE _____

A B C D

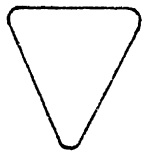
PROBLEM 1 UPPER
As a driver, when you reach a traffic sign of this shape at an intersection while driving, you must:



ANSWER 1U (Page 18)
A. Yield to cross traffic.
B. Watch for an on-coming train.
C. Slow and proceed with caution.
D. Stop completely and yield.

1U

PROBLEM 2 UPPER
What is the meaning of this sign which is triangular with one point downward?



ANSWER 2U (Page 19)
A. Regulatory.
B. Warning.
C. Yield.
D. Railroad crossing.

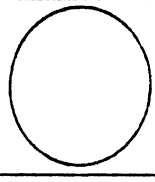
2U

PROBLEM 3 UPPER
At a YIELD sign you must slow to a speed which will enable you to stop quickly, if necessary. You must give the right-of-way to:

ANSWER 3U (Page 19)
A. All of the following.
B. All pedestrians.
C. Other traffic in the intersection.
D. All approaching traffic.

3U

PROBLEM 4 UPPER
This circular sign with a yellow background gives advance warning of:



ANSWER 4U (Page 19)
A. A danger zone.
B. A railroad crossing.
C. A school zone.
D. An intersection.

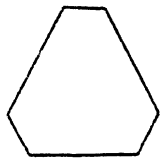
4U

PROBLEM 5 UPPER
At a railroad crossing, when a train is approaching and you must stop, you must do so not less than:

ANSWER 5U (Page 19)
A. 5 ft. from the near rail.
B. No limit established.
C. 10 ft. from the near rail.
D. 15 ft. from the near rail.

5U

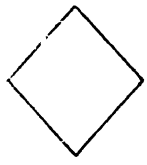
PROBLEM 6 UPPER
A sign of this shape on the rear of a vehicle indicates that it cannot travel faster than 25 mph.



ANSWER 6U (Page 20)
A. True. B. False.

6U

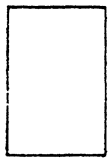
PROBLEM 7 UPPER
This diamond shaped sign may bear symbols or a written message to indicate:



ANSWER 7U (Page 21)
A. A school zone.
B. Some form of hazard ahead.
C. A NO PASSING zone.
D. A pedestrian crossing.

7U

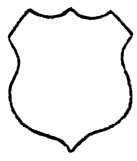
PROBLEM 8 UPPER
A rectangular sign that is taller than it is wide is which classification of sign?



ANSWER 8U (Page 24)
A. Railroad crossing sign.
B. Warning sign.
C. Guide.
D. Regulatory.

8U

PROBLEM 9 UPPER
Any sign which has the shape of a shield is always posted to act as a:



ANSWER 9U (Page 26)
A. Danger warning.
B. Route marker.
C. Animal crossing.
D. School zone.

9U

PROBLEM 10 UPPER
When approaching a traffic control signal displaying a steady red light you must stop behind a crosswalk or stop line and:

ANSWER 10U (Page 27)
A. Wait for a green signal.
B. Proceed with caution.
C. Yield, if necessary.
D. Then proceed when safe.

10U

AP-4U

INSTRUCTIONS TO NON-VOLUNTEERS WHO WERE MAILED THE SAME PACKAGE AS THAT RECEIVED BY THE "WEDMK" GROUP

Inter Office Communication

August 21, 1972

Dear Ford Employe:

Although you are not a participant in the Employe Skilled Driving Program, we thought you would be interested in receiving a copy of the latest Michigan Driver Manual, "What Every Driver Must Know." This manual provides pertinent information relative to laws and safety rules of the road. Anyone who drives in Michigan will find this manual helpful.

The manual which is being used in the Employe Skilled Driving Program, has been programmed for self-pace, self-instruction using a set of 80 multiple choice questions. The answer code is provided below.

Should you have any thoughts or comments concerning the manual or the Program, we would be pleased to hear from you.

Phil Gram
Employe Skilled Driving Program

ANSWER CODE - "WHAT EVERY DRIVER MUST KNOW"

X - 1M Upper		X - 1M Lower		X - 2M Upper		X - 2M Lower	
1 - D	6 - A	1 - B	6 - A	1 - C	6 - D	1 - D	6 - C
2 - C	7 - B	2 - D	7 - C	2 - B	7 - A	2 - B	7 - A
3 - A	8 - D	3 - D	8 - D	3 - B	8 - B	3 - C	8 - A
4 - B	9 - B	4 - B	9 - C	4 - C	9 - D	4 - A	9 - D
5 - D	10 - A	5 - C	10 - D	5 - A	10 - A	5 - C	10 - B

X - 3M Upper		X - 3M - Lower		X - 4M Upper		X - 4M Lower	
1 - B	6 - A	1 - C	6 - B	1 - D	6 - C	1 - B	6 - A
2 - D	7 - C	2 - A	7 - D	2 - B	7 - A	2 - B	7 - B
3 - D	8 - D	3 - B	8 - B	3 - C	8 - A	3 - C	8 - D
4 - B	9 - C	4 - D	9 - A	4 - A	9 - D	4 - A	9 - A
5 - C	10 - D	5 - A	10 - C	5 - C	10 - B	5 - D	10 - A

LISTING OF INDIVIDUAL STUDY MATERIALS AVAILABLE IN THE AUDIO-VISUAL LIBRARY

TITLE	SUBJECT	CONTENTS
PASSING MANEUVERS	"Basic Passing" "Hazard on the Side" "Oncoming Traffic" "Being Passed" "Passing Emergencies"	*Five 35mm color filmstrips, three 12" records and one set instructor guides & driver workbooks*
	Driver Workbooks	
	Instructor Guides	
INTERSECTION MANEUVERS	"Basic Intersection Maneuver" "Through Signal" "Right Turns" "Left Turns" "Merging Traffic"	*Five 35mm color filmstrips, three 12" records and one set instructor guides & driver workbooks*
	Driver Workbooks	
	Instructor Guides	
FREEWAY MANEUVERS	"Entering the Freeway" "Driving at Freeway Speeds" "Judgment Time Driving" "Passing on the Freeway" "Leaving the Freeway"	*Five 35mm color filmstrips, three 12" records and one set instructor guides & driver workbooks*
	Driver Workbooks	
	Instructor Guides	
EMERGENCY MANEUVERS	"Emergency Braking Skills" "Controlling Skids" "Wheels Off Pavement" "Sudden Loss of Vision"	*Four 35mm color filmstrips, two 12" records and one set instructor guides & driver workbooks*
	Driver Workbooks	
	Instructor Guides	
DRIVING STRATEGY	"Decision Pattern" "Strategic Positioning" "Adjusting to the Changing Scene" "Critical Maneuvers - Skids" "Emergency Problems"	*Five 35mm color filmstrips, three 12" records and one set instructor guides & driver workbooks*
	Driver Workbooks	
	Instructor Guides	
SEEING HABITS FOR EXPERT DRIVING	A complete unit based on the Smith System	*1 35mm color filmstrip, 1 12" record and 1 27 page Companion Workbook*
	Companion Workbooks	
9 OUT OF 10	A self appraisal approach to driving a look at the other driver	*1 35mm color filmstrip 1 12" record and 5 Companion Workbooks*
	Companion Workbooks	

Ford Employe Skilled Driving Program

AUDIO - VISUAL LIBRARY
MEMBERSHIP CARD

Library Membership Card

This is to certify that

_____

is a duly registered member of this life-saving Program and therefore entitled to special rights and privileges of the Audio-Visual Library facilities located at World Headquarters, Main Lobby - East, Dearborn, Michigan

Date _____

No _____

_____
Phil Gram
Program Manager

APPENDIX 2

BIOGRAPHICAL AND DRIVING HISTORY QUESTIONNAIRE

EMPLOYE SKILLED DRIVING PROGRAM
BIOGRAPHICAL AND DRIVING HISTORY INFORMATION

We are pleased that you have volunteered to participate in the FORD Skilled Driving Program. Every effort will be made to make this a worthwhile and enjoyable experience. Information regarding your driving experience will be used by the University of Michigan to statistically evaluate various aspects of the program.

Please fill out the following form:

		Punch Card Code
1.	Name: _____	
2-5.	# _____ / 2 0 / _ / _ / _ / (Leave Blank)	1 - 10
6.	Date: (Month-Day-Year) _ _ / _ _ / _ _	11 - 16
7.	How often do you normally drive a Company owned car on Company business? (Pool, lease, engineering, executive, etc.)	
	(circle)	
	Seldom if any 1	
	Once or twice per month 2	
	Once or twice per week 3	17
	Nearly every day 4	
	Several hours a day 5	
--	Please list tag numbers of Company vehicles you usually drive (if any).	
8.	Tag number _ _ _ _ _ _ _	18 - 24
9.	Tag number _ _ _ _ _ _ _	25 - 31
10.	How many miles is your daily trip to and from work? _ _ _	32 - 35
11.	In total, how many thousand miles a year do you average? _ _	36 - 37
12.	How many years have you been driving? _ _	38 - 39
--	Vehicles used in this program are 1972 Ford Galaxies with automatic transmission, power steering and power brakes.	(40 blank)
13.	Are you used to driving this kind of car with this kind of equipment? Yes __ (1) No __ (0)	41
--	If your answer to question 13 was no, please check any of the following: (Questions 14, 15 and 16)	
14.	Drive smaller car __ (1) Drive larger car __ (2)	42

		Punch Card Code
15.	Not used to power steering __ (1)	43
16.	Not used to power brakes __ (1)	44
17.	What size vehicle do you drive most frequently?	
	(circle) Small (Pinto) 1 Compact (Maverick) 2 Intermediate (Torino) 3 Standard (Ford) 4 Luxury (Mercury-Lincoln) 5	45
18.	Have you ever taken a high school driver education course? Yes __ (1) No __ (0)	46
19.	Did you learn to drive in a formal driving school other than a high school driver education course? Yes __ (1) No __ (0)	47
20.	Have you taken other driver improvement courses? Yes __ (1) No __ (0)	48
21.	Do you feel that it is worthwhile for Ford Motor Company to provide skilled driving instruction for its employes? Yes __ (1) No __ (0)	49
22.	Please give your reasons for this answer:	50 - 51
	_____	
	_____	
23.	Why did you decide to take advantage of this course?	52 - 53
	_____	
	_____	
24.	At some future date, may we check your driving record (accidents and violations) with the understanding that it will be used for statistical group analysis only and will be held in confidence? Yes __ (1) No __ (0)	54
--	Please list your:	
25.	Michigan driver's license expiration date __/__/__	55 - 60
26.	Michigan driver's license number _/___/___/___/___	61 - 73
27.	How many members of your immediate household, including yourself, are drivers? __	74

APPENDIX 3

COURSE EVALUATION SURVEY MAILED TO PARTICIPANTS FOUR MONTHS AFTER POST-TEST

Ford Motor Company

The American Road
Dearborn, Michigan 48121
February, 1973

Dear Ford Employe:

We are pleased that you are a participant in the Employe Skilled Driving Program. The University of Michigan's Highway Safety Research Institute is conducting an independent evaluation of the Program; to assist them, we have been asked to mail the enclosed questionnaire to all participants, whether or not they completed their respective assigned courses.

Your opinion of the program is very important to HSRI's evaluation. Would you, therefore, please take a few minutes to complete the questionnaire and return it to me, either by Company or regular mail:

Phil Gram, Manager
Employe Skilled Driving Program
Ford Motor Company
World Headquarters - Lobby - East
Dearborn, Michigan 48121

Information obtained from this survey is confidential. After the information is analyzed by HSRI, results will be presented to Ford management in summary form only. A code number on each form is used to enable a follow-up reminder to those who might forget to return their questionnaires in a reasonable time.

Thanks for your cooperation on behalf of HSRI and the Employe Skilled Driving Program.

Phil Gram, Manager
Employe Skilled Driving Program

Enclosure

HIGHWAY SAFETY RESEARCH INSTITUTE
UNIVERSITY OF MICHIGAN

FORD EMPLOYE SKILLED DRIVING PROGRAM

Your responses to the following questions about the Ford Employe Skilled Driving Program, in which you participated recently, would be of considerable assistance to us:

1. How useful was the Skilled Driving Program to you?

(1) ____ Outstandingly Useful
(2) ____ Very Useful
(3) ____ Moderately Useful
(4) ____ Not Very Useful
(5) ____ Not at All Useful

2. What was the most useful part of the program?

_____

_____

3. What was the least useful part of the program?

_____

_____

4. Can you give an example from your driving experience since the end of the program:

(a) Of a way it was helpful to you? _____

_____

_____

(b) Of a way it confused or mislead you? _____

_____

_____

5. How do you feel about the Company providing a skilled driving program for employes?

(1) ____ I strongly disapprove
(2) ____ I tend to disapprove
(3) ____ I don't care
(4) ____ I tend to approve
(5) ____ I strongly approve

6. Why did you decide to take advantage of this course?

_____

_____

_____

7. How would you rate your driving ability before the Skilled Driving Program?

(1) ____ Excellent
(2) ____ Above Average
(3) ____ Average
(4) ____ Below Average
(5) ____ Poor

How would you rate it after the Skilled Driving Program?

(1) ____ Excellent
(2) ____ Above Average
(3) ____ Average
(4) ____ Below Average
(5) ____ Poor

8. What suggestions do you have for changes or additions to the Program?

Program content _____

_____

Instructors _____

_____

Teaching Methods _____

_____

Instructional Materials (films, test cards) _____

_____

9. How often would you like to take a driving program such as this one?

(1) ____ Not again
(2) ____ Every two years
(3) ____ Every three years
(4) ____ Every four years
(5) ____ _____

10. Out of the following three groups of drivers, which group do you think is responsible for the highest percentage of all traffic accidents:

(1) ____ Drivers who have already had a serious accident.

(2) ____ Drivers with grossly anti-social behavior

(3) ____ "Average" drivers.

11. Would you like to have this program made available to members of your family? (1) ____ Yes (2) ____ No

12. Do you have any other comments about the Ford Employe Skilled Driving Program?

_____

_____

_____

_____

_____

_____

Return to:

Phil Gram, Manager
Employe Skilled Driving Program
Ford Motor Company
World Headquarters - Lobby - East
Dearborn, Michigan 48121

Code:_____

APPENDIX 4

WRITTEN TESTS AND NEAR-MISS REPORT CARD

INTRODUCTION TO POST TEST

You will have a total of 15 minutes to complete the attached four-part Post Test.

Pace yourself. Work quickly and accurately. Most of the questions are multiple choice and require only one answer per question.

Your instructor will tell you when to begin the test. DO NOT TURN THIS PAGE UNTIL THEN.

_____
(Name - Please Print)

Check Appropriate Box:

COURSE ASSIGNMENT

☐ What Every Driver Must Know

☐ Self-Teach Course

☐ 4-Hr. ☐ 8-Hr. ☐ *N.A.

*Not Assigned

Do Not Write In this Space

I.D. Number _ _ _ _ _	(1-5)
Data Type 6 0	(6-7)
Instructor Code _	(8)
Time of Day _	(9)
Section Number _	(10)
Date _ _ _ _ _ _ (Mo. Day Yr.)	(11-16)

POST TEST
FORD SKILLED DRIVING PROGRAM

Part I

General Driving Knowledge (Check only one answer per question)

1. When changing lanes you should:

   ☐ a. Signal, look behind you and check your mirror

   ☐ b. Check your mirror, look behind you and signal

   ☐ c. Look behind you, signal and check your mirror

   ☐ d. Signal, check your mirror and look behind you

2. If a speed limit is not posted before a curve:

   ☐ a. Continue at the same speed

   ☐ b. Assume that it is better to take this curve at a higher speed

   ☐ c. Slow to 35 - 40 mph

   ☐ d. Judge how sharp the curve is and change your speed accordingly

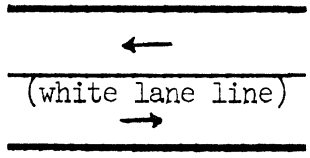
3. The solid, white lane line means:

   ☐ a. Changing lanes permitted only for turning vehicles

   ☐ b. Trucks may not change lanes; other vehicles are permitted to do so with care

   ☐ c. Changing lanes is not allowed

   ☐ d. Changing lanes is discouraged



4. If the signal at a railroad crossing does not indicate that a train is coming you should:

   ☐ a. Speed up and cross the tracks quickly

   ☐ b. Continue at the same speed and check for a train before crossing

   ☐ c. Slow down and look both ways

   ☐ d. Come to a complete stop before continuing across

5. Hydroplaning, a situation where the vehicle's tires lose contact with the pavement at high speeds, is most likely to occur:

- ☐ a. During a heavy rain
- ☐ b. During a gentle rain
- ☐ c. During a light mist or fog
- ☐ d. Only on ice

6. The Basic Speed Law requires that you must drive a vehicle on a highway at a speed that is:

- ☐ a. Always at the posted speed limit
- ☐ b. Slower than other traffic
- ☐ c. Steady and stabilized
- ☐ d. Careful and prudent

7. The responsibility for maintaining a vehicle in safe operating condition at all times rests with:

- ☐ a. The owner of the vehicle
- ☐ b. The mechanic
- ☐ c. The driver of the vehicle
- ☐ d. The manufacturer

8. When leaving the freeway, you should reduce your speed to meet the new driving situation. The proper place to adjust your speed is:

- ☐ a. In Exit Lane before Exit Ramp
- ☐ b. On the Exit Ramp
- ☐ c. In right hand travel lane
- ☐ d. At the end of the Exit Ramp

Part I

9. You are permitted to leave the pavement to pass a vehicle on the right while it is making a left turn:

- ☐ True
- ☐ False

10. When pedestrians appear at crosswalks, drivers must reduce speed or stop:

- ☐ True
- ☐ False

POST TEST
FORD SKILLED DRIVING PROGRAM

Part II

The Decision Pattern in Driving Maneuvers (Check only one answer per question)

1. The six steps of the decision pattern are as follows:

(a) Initial Alert
(b) Point of Decision
(c) Scan the Scene
(d) Check Conflict Areas
(e) Point of No Return
(f) Initial Move

- ☐ a. (a), (c), (d), (f), (e), (b)
- ☐ b. (a), (c), (b), (d), (f), (e)
- ☐ c. (a), (c), (f), (d), (b), (e)
- ☐ d. (a), (f), (c), (d), (e), (b)

2. The decision pattern can be used in the following traffic maneuvers:

- ☐ a. All maneuvers except on freeways
- ☐ b. Almost all traffic maneuvers
- ☐ c. Very few traffic maneuvers
- ☐ d. Only during left turns, right turns and changing lanes

3. If you have another alert in the middle of a decision pattern you should immediately:

- ☐ a. Abort the first pattern and start again
- ☐ b. Reevaluate to determine priorities
- ☐ c. Overlap the two patterns
- ☐ d. Complete the first pattern before starting the second

4. Which of the following is not a method of communicating with other drivers when you make a right turn?

- ☐ a. Tap your brakes
- ☐ b. Flash your right turn signal
- ☐ c. Position your vehicle in the proper lane
- ☐ d. Stop for the traffic signal to change to green

Part II

5. Which of the following would most likely cause the decision pattern to be foreshortened?

   ☐ a. An emergency

   ☐ b. Night driving

   ☐ c. When there is no traffic

   ☐ d. All of the above

6. Which of the following is the least desirable way to influence the actions of a driver who is tailgating?

   ☐ a. Flash brake lights

   ☐ b. Turn on your lights

   ☐ c. Slow down quickly

   ☐ d. Let him pass

POST TEST
FORD SKILLED DRIVING PROGRAM

Part III

Strategic Positioning (Check only one answer per question)

Strategic positioning is defined as keeping the best possible space cushion around your car to protect your margin of safety.

1. Which of the following will not help your strategic positioning in traffic?

   ☐ a. Balanced spacing front, rear and to the sides

   ☐ b. Adjusting your speed to that of the traffic around you

   ☐ c. Staying in the right lane as much as possible

   ☐ d. Changing lanes as required for good sight distance

2. One of the following vehicle controls is used in strategic positioning:

   ☐ a. The headlights

   ☐ b. The shift lever

   ☐ c. Parking brake

   ☐ d. Accelerator

3. Strategic positioning allows you more time and space to make decisions and execute maneuvers. In which of the following situations is strategic positioning most important?

   ☐ a. In an emergency

   ☐ b. In a parallel parking maneuver

   ☐ c. When you are entering a freeway

   ☐ d. During a left turn

4. Which of the following is not a good reason for changing lanes on a freeway?

   ☐ a. To pass a slower vehicle

   ☐ b. To improve your sight distance

   ☐ c. To slow down a speeder

   ☐ d. To improve your space cushion

Part III

5. When on an Entrance Ramp to a freeway with a vehicle in front of you and one behind you:

   ☐ a. Try to have more space between you and the vehicle behind

   ☐ b. Try to have more space between you and the vehicle in front

   ☐ c. Stay close to the vehicle in front

   ☐ d. Divide the space equally between the vehicle in front and the one behind

6. Which three vehicle controls are used to position your vehicle in traffic?

   ☐ a. Turn signals, mirrors, brake lights

   ☐ b. Brake, mirrors, turn signals

   ☐ c. Accelerator, brake, steering wheel

   ☐ d. Outside mirror, turn signal, shift lever

7. Which of the following rationale is most important for changing lanes on a freeway?

   ☐ a. To improve sight distance and spacing

   ☐ b. To pass or permit others to pass you

   ☐ c. To improve spacing and go a little slower than traffic

   ☐ d. To go a little faster than traffic and have better visibility

POST TEST
FORD SKILLED DRIVING PROGRAM

Part IV

Emergencies and Other Driving Problems (Check only one answer per question)

1. Your car has automatic transmission, power steering and power brakes. You are turning right at an intersection and your engine stalls, which of the following happens?

   ☐ a. Steering effort will be greatly increased

   ☐ b. You cannot shift

   ☐ c. The car will not coast unless you put it in neutral

   ☐ d. The steering wheel will lock

2. Which of the following may result in a sudden loss of braking control?

   ☐ a. Driving through a deep puddle of water

   ☐ b. Having a soft tire

   ☐ c. Having the accelerator stick

   ☐ d. Driving at high altitudes

3. Which one of the following will not help the driver who has a sudden loss of visibility?

   ☐ a. Point of no return

   ☐ b. Better sight distance

   ☐ c. Strategic positioning

   ☐ d. Good communications

4. Your windshield is suddenly splattered by muddy water. In this emergency your first problem to solve is:

   ☐ a. Regaining your lost vision

   ☐ b. Getting your wheels back on the pavement

   ☐ c. Warning any other drivers behind you

   ☐ d. Regaining your speed control

Part IV

5. To properly regain positioning control after your wheels have gone off the pavement, you must resist the immediate urge to:

   ☐ a. Continue steering straight ahead

   ☐ b. Whip your wheels back onto the pavement

   ☐ c. Accelerate for better steering control

   ☐ d. Pull farther onto the road shoulder

6. What is the recommended method to warn drivers behind you of an emergency?

   ☐ a. Blow horn

   ☐ b. Flash emergency signal

   ☐ c. Flash lights

   ☐ d. Pump brakes lightly

7. Do you consider sudden braking or swerving to avoid hitting the car ahead an emergency maneuver?

   ☐ a. Yes

   ☐ b. Most of the time

   ☐ c. Occasionally

   ☐ d. No

8. You were asked to keep a scorecard of your near accidents. Please fill in the appropriate answers:

   ☐☐ a. Number of NEAR ACCIDENTS recorded

   ☐☐ b. Number of weeks since start of recording

   ☐ c. Kept score but card not available

   ☐ d. Did not keep a record

DRIVER INSTRUCTIONS

— This card will help you remember near-accident situations.
— Please fill in the date you begin your observations (top line).
— When a near-accident occurs, fill in date, time, approximate location and code for each situation as soon as possible.
— Include all situations, even if the other driver was at fault.
— Continue observations for several months to see if the number of near-accident situations are decreasing or increasing.
— Estimate total number of hours you drove between "start and end of observations."

American Road Skilled Driving Programs
World Headquarters, Dearborn, Michigan 48121

NEAR ACCIDENT SCORECARD

	DATE	TIME	LOCATION	CODE
—	/ /72	*Start of observations*		
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
—	/ /72	*End of observations*		

CODE

BH — Braked Hard
S — Swerved
AR — Accelerated Rapidly
OC — Other Car — braked, swerved or accelerated
NM — Near Miss — no evasive action

Approximate number of driving hours between start & end of observations:
[ ] hours

For Instructions See Other Side

APPENDIX 5

UNUSUAL USES TEST SHEETS

You are driving along a two lane country road at night behind a farm truck going 20 mph. You want to pass. List all the things you can do to analyse the conditions and maintain this as a safe-driving situation.

You are driving on a two lane road during daylight hours at 35-40 mph. You notice in your rear view mirror a car following closely is starting to pull out to pass you. List all the things you can do to analyse the conditions and maintain this as a safe-driving situation.

APPENDIX 6

DESCRIPTION OF OBJECTIVE BEHIND-THE-WHEEL VARIABLES AND DATA SHEET

OBJECTIVE ROAD TEST

. Description of Data Sheet Variables

Data output provided by the computer tapes include information on steering wheel reversal rate, vehicle speed, brake applications, total run time, delay time, and an index number rating system.

Scoring sheet items #2 thru #7 relate to the average rate of steering wheel movement at the six periods of observation, "a" thru "f". These are followed on the sheet by an arithmetic mean of steering wheel reversal rates and a ratio of reversal rate during the freeway stress versus the freeway pre-stress periods.

Items #10 and #11 are measures of total time and delay time, respectively. The latter is defined as the times at which vehicle speed drops below 20 M.P.H. Items #12 thru #14 are index numbers relating to the overall measure of composite driver behavior, while the next item is another ratio dealing with freeway stress versus freeway pre-stress behavior.

Item #16 is a measure of the number of brake applications, while items #17 thru #20 are measures of vehicle speed: #17 being the maximum speed achieved during the run and #18 thru #20 indicating the average speed during the three freeway periods of observation. Items #21 and #22 are ratios of, respectively, speed and steering reversal rates during different observation periods.

EMPLOYE SKILLED DRIVING PROGRAM

DATA SHEET FOR OBJECTIVE DRIVER EVALUATION

Variable No.				Frames	Punch Card Code
	Name: _____				
	#_ _ _ _ _ _ _ _ _ _			(10)	1 - 10
	Date: (Month-Day-Year) _ _ _ _ _ _			(6)	11 - 16
1	SRRB		_ _ _	(3)	17 - 19
2	SRRa		_ _	(2)	20 - 21
3	SRRb		_ _	(2)	22 - 23
4	SRRc		_ _	(2)	24 - 25
5	SRRd		_ _	(2)	26 - 27
6	SRRe		_ _	(2)	28 - 29
7	SRRf		_ _	(2)	30 - 31
8	SRRA	Add (2) to (7)	_ _ _	(3)	32 - 34
9	$SRRR^1$	$\frac{(6)}{(2)}$	_ _ _	(3)	35 - 37
10	TT		_ _ _	(3)	38 - 40
11	DT		_ _ _	(3)	41 - 43
12	INa		_ _ _ _	(4)	44 - 47
13	INe		_ _ _ _	(4)	48 - 51
14	INf		_ _ _ _	(4)	52 - 55
15	INR	$\frac{(13)}{(12)}$	_ _ _	(3)	56 - 58
16	BA		_ _ _	(3)	59 - 61
17	MSP		_ _ _	(3)	62 - 64
18	ASPa		_ _ _	(3)	65 - 67
19	ASPe		_ _ _	(3)	68 - 70
20	ASPf		_ _ _	(3)	71 - 73
21	ASPR	$\frac{(c)}{(f)}$	_ _ _	(3)	74 - 76
22	$SRRR^2$	$\frac{(4)}{(3)}$	_ _ _	(3)	77 - 79

APPENDIX 7

SUBJECTIVE BEHIND-THE-WHEEL TEST: TEST FORM, EXPANDED VERSION OF ITEMS, ITEM ANALYSES, AND STUDY ON THE RELATIONSHIP OF THE TEST TO MAJOR BIOGRAPHICAL VARIABLES

Employe Name _____

ID Number __ __ __ __ __ (1-5)

Data Type 3 __ (6-7)

Instructor Code __ (8)

Time of Day __ (9)

Section Number __ (10)

Date __ __ __ __ __ __ (11-16)
(Month, Day, Year)

FREEWAY SEQUENCE - PRE STRESS

1. Kept scanning far ahead? []Yes 1 []No 0 (17)
2. Habitual mirror and head checks? []Yes 1 []No 0 (18)
3. Deliberate attempts to improve cushion? []Yes 1 []No 0 (19)
4. Average gap time at speeds over 45 mph? []½ 1 []1 2 []1½ 3 []2 4 []2½ 5 []3 6 []3+ 7 (20)

Was traffic []unusually light? 1 []about normal? 2 []unusually heavy? 3 (21)

Was road surface []Dry 1 (Note exceptional []Slippery 2 conditions) (22)

NOTES

LEFT TURN SEQUENCE - PRE STRESS

5. Full scan before lane change? []Yes 1 []No 0 (23)
6. Lane change timing wise? []Yes 1 []No 0 (24)
7. Adequate signalling in lane change and left turn? []Yes 1 []No 0 (25)
8. Checked conflicts ahead/behind to sides, on thru street? []Yes 1 []No 0 (26)
9. Checked obstructions in street being turned into? []Yes 1 []No 0 (27)
10. Wheels turned straight ahead? []Yes 1 []No 0 (28)

Was traffic []unusually light? 1 []about normal? 2 []unusually heavy? 3 (29)

Was road surface []Dry 1 (Note exceptional []Slippery 2 conditions) (30)

NOTES

LEFT TURN SEQUENCE - STRESS

5. Full scan before lane change? []Yes 1 []No 0 (31)
6. Lane change timing wise? []Yes 1 []No 0 (32)
7. Adequate signalling in lane change and left turn? []Yes 1 []No 0 (33)
8. Checked conflicts ahead/behind/ to sides, or thru street? []Yes 1 []No 0 (34)
9. Checked obstructions in street being turned into? []Yes 1 []No 0 (35)
10. Wheels turned straight ahead? []Yes 1 []No 0 (36)

Ability to maintain secondary task []excellent 1 []reasonable 2 []poor 3 (37)

Was traffic []unusually light? 1 []about normal? 2 []unusually heavy? 3 (38)

Was road surface []Dry 1 (Note exceptional []Slippery 2 conditions) (39)

NOTES

LEFT TURN SEQUENCE - RECOVERY

5. Full scan before lane change? []Yes 1 []No 0 (40)
6. Lane change timing wise? []Yes 1 []No 0 (41)
7. Adequate signalling in lane change and left turn? []Yes 1 []No 0 (42)
8. Checked conflicts ahead/behind/ to sides, on thru street? []Yes 1 []No 0 (43)
9. Checked obstructions in street being turned into? []Yes 1 []No 0 (44)
10. Wheels turned straight ahead? []Yes 1 []No 1 (45)

Was traffic []unusually light? 1 []about normal? 2 []unusually heavy? 3 (46)

Was road surface []Dry 1 (Note exceptional []Slippery 2 conditions) (47)

NOTES

Employe Skilled Driving Program

HIGHWAY SAFETY RESEARCH INSTITUTE
Huron Parkway and Baxter Road
Ann Arbor, Michigan 48105
Telephone (313) 764-0748

THE UNIVERSITY OF MICHIGAN

FREEWAY SEQUENCE - STRESS

1. Kept scanning far ahead? []Yes 1 []No 0 (48)
2. Habitual mirror and head checks? []Yes 1 []No 0 (49)
3. Deliberate attempts to improve cushion? []Yes 1 []No 0 (50)
4. Average time gap at speeds over 45 mph? []½ 1 []1 2 []1½ 3 []2 4 []2½ 5 []3 6 []3+ 7 (51)

Ability to maintain secondary task []excellent 1 []reasonable 2 []poor 3 (52)

Was traffic []unusually light? 1 []about normal? 2 []unusually heavy? 3 (53)

Was road surface []Dry 1 (Note exceptional []Slippery 2 conditions) (54)

NOTES

FREEWAY SEQUENCE - RECOVERY

1. Kept scanning far ahead? []Yes 1 []No 0 (55)
2. Habitual mirror and head checks? []Yes 1 []No 0 (56)
3. Deliberate attempts to improve cushion? []Yes 1 []No 0 (57)
4. Average gap time at speeds over 45 mph? []½ 1 []1 2 []1½ 3 []2 4 []2½ 5 []3 6 []3+ 7 (58)

Was traffic []unusually light? 1 []about normal? 2 []unusually heavy? 3 (59)

Was road surface []Dry 1 (Note exceptional []Slippery 2 conditions) (60)

NOTES

ADDITIONAL DATA

A. Did driver complain about having to wear seat belt? []Yes 1 []No 0 (61)
B. Did driver wear shoulder belt? []Yes 1 []No 0 (62)
C. Unusual physical or mental condition? []Yes 1 []No 0 (63)
Specify _____

SUBJECTIVE ROAD TEST:

Detailed Instructions for Items (Key Words in Parentheses)

Freeway Sequence:

1. Did driver scan far enough ahead all the time? If there is any evidence that driver is "surprised" by traffic events--e.g., having to brake sharply to maintain headway, or responding late to direction signs--mark "No". Also mark "No" if driver allows vision to be blocked by traffic in front. (Kept scanning far ahead?)

☐ Yes
☐ No

2. Did driver habitually check potential conflict points to both the sides and rear? Take special note of routine checks of on-ramps and off-ramps and merging situations, regardless of the lane he is in. If there is no attempt to make either a head check or a several-angled mirror check in potential conflict situations, mark "No". (Habitual mirror and head checks?)

☐ Yes
☐ No

3. Did driver preserve his space cushion, or improve it by a change in position and speed? If, during an observational sequence, a driver both deliberately improves his position and allows it to worsen, mark "Yes" or "No" according to whether his "good" actions were more significant than his "bad" actions. If no opportunities to score this arise, check neither box and write "N.A." on the form. Note that the keywords that follow do not suggest all of the above, but I don't want to change the forms at this point. (Deliberate attempts to improve cushion?)

☐ Yes
☐ No

4. What is the average headway driver is prepared to maintain at speeds over 45 mph? "Weight" the average towards a shorter time if drivers seems so inclined, but was only briefly observed with a short gap. (Average gap time at speeds over 45 mph?)

☐ ½ sec.
☐ 1 sec.
☐ 1½ sec.
☐ 2 sec.
☐ 2½ sec.
☐ 3 sec.
☐ 3+ sec.

Left Turn Sequence:

5. Before making lane change did driver check ahead, behind and sides? The blind spot must be checked: if driver did not make a head check or look in his side mirror from several angles, mark "No". (Full scan before lane change?)

☐ Yes
☐ No

6. Was driver's decision to complete his lane change timed wisely, such that he minimized hazard and congestion? "No's" include: changing unnecessarily late; changing in major intersections where there is waiting traffic unless the

alternatives were more hazardous; changing unnecessarily early, such that you become trapped by vehicles turning left at previous intersections (again, becoming so trapped will not always mean that the driver failed to choose the most advantageous time to change--you will have to compare your assessment of the situation against his). (Lane change timing wise?)

☐Yes
☐No

7. Did driver signal adequate information--horn, turn signals, eye contact etc.--in both the lane change and left turn maneuvers? Turn signals should be used for all lane changes and turns on these busy urban streets. Remember that the timing of signals is important. Mark "No" if he changes lanes before several clicks of the turn signal--but be lenient in the case of the occasional rapid evasive lane change, providing adequate visual checks were made, and he did not bring such a critical situation upon himself. (Adequate signalling in lane change and left turn?)

☐Yes
☐No

8. Did driver adequately check traffic ahead, behind, and to the sides on the street from which he makes his turn? This refers to his actions after the lane change. (Checked conflicts ahead/behind/to sides on through street?)

☐Yes
☐No

9. Did driver adequately check street he is turning into for potential obstructions which might leave him "stranded" (pedestrians, cars, etc.) including checking sidewalks? (Checked obstructions in street being turned into?)

☐Yes
☐No

10. Did driver keep his wheels straight ahead before turn? Give him about 15 degrees for "arm-droop", as this is unlikely to change the trajectory too much if the car is struck. Remember, he is less vulnerable if other cars are waiting behind him. Also, mark "No" if he turns the wheels before he starts his turn, but he can begin to move the wheel as he is passed by the last oncoming car before an acceptable gap. (Wheels straight ahead before turn?) "N.A." may be used for this item.

☐Yes
☐No

SUBJECTIVE BEHIND-THE-WHEEL
ITEM ANALYSES

Inter-Observer Reliability

	November 72		March 73		May 73	
Item	*N*	γ	*N*	γ	*N*	γ
1	24	0.75	27	0.52	33	0.64
2	24	0.42	27	0.33	33	0.40
3	21	0.24	22	0.05	33	0.27
4	23	0.75	27	0.93	33	0.88
5	24	0.42	27	0.11	33	-0.03
6	23	0.29	27	0.46	33	0.64
7	24	0.06	27	0.48	33	0.03
8	24	-0.08	27	0.15	33	0.21
9	23	0.37	27	0.41	33	0.27
10	24	0.42	27	0.30	28	0.47

$$\gamma_i = \frac{A_i - D_i}{N_i}$$

A_i = Number of agreements between observations on item i.
D_i = Number of disagreements between observations on item i.
N_i = Total number of observations on item i.

Item Difficulty

	November 72		March 73		May 73	
Item	*N*	*P*	*N*	*P*	*N*	*P*
1	21	0.95	21	1.00	27	0.93
2	17	0.47	18	0.28	23	0.61
3	13	0.62	11	0.45	21	0.62
4	n/a	n/a	n/a	n/a	n/a	n/a
5	18	0.44	14	0.43	16	0.63
6	15	0.80	20	0.80	28	0.89
7	17	0.12	20	0.50	17	0.71
8	11	0.36	15	0.40	20	0.50
9	15	0.27	19	0.63	21	0.86
10	17	0.88	16	0.94	21	0.90

$$P_i = \frac{S_i}{N_i}$$

S_i = Number of agreed observations of correct execution of item i.
N_i = Total number of agreed observations of item i.

THE INFLUENCE OF SOCIO-ECONOMIC STATUS, AGE AND DRIVING EXPERIENCE ON A TEST OF SELECTED DRIVING ABILITIES

Martin E. Lee

April 1973

(Abridged, October 1973)

Abstract

The writer has previously developed a test of selected automobile driving abilities as part of the evaluation of an adult driver improvement program. Such tests are frequently suspected to be dependent upon certain biographical variables, most commonly socio-economic status, age, and driving experience. This study explores the relationship between pre-treatment test scores and these three variables (N=508). Several significant regression equations were obtained, but even with all three independent variables in the model, only 4.4% of the test score variance could be explained. This result was tested and upheld using a split half technique.

1.0 Introduction

In recent years, the authors of tests of automobile driving ability have been criticized in the literature for failing to establish the degree of dependence between test scores and certain biographical variables. Most often mentioned are socio-economic status, age, and driving ability. The purpose of this study is to explore relationships between scores on the SBTW test and these three biographical variables.

2.0 Analytical Procedures

The analysis was performed in three phases: preliminary exploration; development and verification of the regression model; and a split-half reliability check of the model.

2.1 Preliminary Exploration

This consisted of the selection of dependent and independent variables, and initial searches of the data.

2.1.1 Selection of the dependent variable. The SBTW test score (Y) consists of 3 sets of 9 binary observations taken over fixed stretches of highway. The maximum score is 27 (merit points); the minimum is 0. The items relate closely to a sample of behaviors which rationally should change as a result of the training courses.

There are three applications of this test: a pre-testing, post-testing, and follow-up testing. Intuitively, the pre-test score was preferable, because to that point all subjects had been treated alike. This included the randomization of treatment groups, and age and salary grade stratifications with the groups, over time of day, day of week, and observer. However, the procedures included partial replication of the analyses described below using post-test scores. The results were similar enough to justify standardizing on pre-

test scores for this study. Follow-up test data was not used as it was incomplete at the time of analysis; and pre-test to post-test difference scores were not attempted as it was felt that their use would raise issues beyond the scope of the present study.

2.1.2 Selection of the independent variables. The initial choice of socio-economic status, age, and driving experience was suggested by criticisms of previous research.

Some early investigations in the HSRI project confirmed that company salary grade was a viable summative variable for socio-economic factors, normally considered in studies of this type for stratified random sampling, such as income, educational level and social class. Salary grade was hence adopted as the first independent variable. Although it is an ordinal scale from 0 to 18, it may be treated as interval with tolerable accuracy.

The second independent variable is age, which was simply transformed from dates of birth derived from company records. It is expressed in one year increments.

Driving experience is a difficult variable to obtain. A major research effort used six variables* to define 26 classes of driving exposure and recommended that all six variables be retained in future studies (Carroll et al., 1971)[2]. Only two variables are available on exposure; the biographical questionnaire asks the number of years of driving experience and approximate number of miles driven per year. The product of these two (expressed in thousands of miles), and years driven were both attempted. The former, an arbitrary estimate of total miles driven, was adopted a priori as the third independent variable as it was less correlated with age than was years driven. Later comparisons showed that total miles was marginally more useful in the regression model.

2.1.3 Initial data searches. A correlation matrix of

---

*The six variables were: age, sex, vehicle type (i.e. auto v. truck, etc), model year, day v. night, and road type.

three candidate Y's and four candidate X_i's was run using MIDAS* to facilitate variable selection. This is reproduced in Table 1. Relationships between Y and the chosen X_i's although low, were enough to justify attempts to build a regression model.

TABLE 1. Correlations between Candidate Dependent and Independent Variables.

N=386 DF=384 R@.950=.0998 R@.990=.1310

VARIABLE	SBTW Pre-Test	SBTW Post-Test	SBTW Post-Pre	Salary Grade	Age	Total Miles
SBTW Pre-Test	1.0000					
SBTW Post-Test	.3171	1.0000				
SBTW Post-Pre	-.4880	.6731	1.0000			
Salary Grade	.0631	.0234	-.0276	1.0000		
Age	-.1588	-.0653	.0637	.2554	1.0000	
Total Miles	.0390	-.0979	-.1205	.2765	.4691	1.0000
Years Driven	-.0767	-.0932	-.0259	.2589	.6811	.7716

It was decided to set a significance level of $\alpha=0.05$ for this purpose. A more stringent level would have been chosen (given the N of 508) but for the degree of measurement of variability inevitable in observational tests such as SBTW. A less stringent level could not be justified as the data will be used to make decisions about future phases of the training program.

2.2 Development and Verification of the Regression Model

A forward stepwise selection of regression was run using MIDAS. An equation was chosen and the residuals were plotted against predicted SBTW pre-test scores.

---

*Michigan Interactive Data Analysis System--a comprehensive statistical software package.

2.3 Split Half Reliability Check of the Regression Model

In order to substantiate the model derived from the entire data set, a split-half reliability technique was used. This was in four stages:

Stage (i): the randomization option in MIDAS:TRANS* was used to assign the 508 cases with valid SBTW pre-test data into two random samples.

Stage (ii): regression equations were generated from both samples, and their residuals plotted against their respective predicted SBTW scores.

Stage (iii): the equation from the first random sample was applied to the second random sample, using MIDAS transformations, to obtain residuals. These were then plotted against the predicted SBTW scores for the second sample.

Stage (iv): the regression equations from the two samples were compared by plotting the second sample values predicted by the first sample equation against the second sample values predicted by the second sample equation. A similar comparison was made between the former, and the second sample values predicted by the equation derived from the entire data set.

3.0 Discussion of Results

3.1 Stepwise Selection of Regression

The MIDAS stepwise procedure yielded three equations, the results of which are summarized in Table 2. These equations are essentially similar in standard error and significance, and R^2 values are so low that it is difficult to make judgements as to which equation is preferable. At the a priori α level of 0.05, equation 3 should be rejected because of the poor significance of the total mileage variable (0.1102). However, it

*TRANS is a sub-program of MIDAS.

would be difficult to argue the exclusion of driving experience from any model designed to reveal the dependent of the SBTW test score, because there is a high expectancy that it would influence a test of driving ability. Therefore, equation 3 was selected. It is:

$$\hat{Y} = 14.828 + 0.17356X_1 - 0.10569X_2 + 0.00141X_3$$

TABLE 2: Summary of Regression Equation Results Obtained by Forward Stepwise Selection (N=508).

EQN #	X_i's INCLUDED	t α FOR X_i's	R^2	S.E.	EQN α
1	Age	0.0004	0.025	4.47	0.0004
2	Salary Grade	0.0063	0.039	4.45	0.0000
	Age	0.0000			
3	Salary Grade	0.0177	0.044	4.44	0.0000
	Age	0.0000			
	Total Miles	0.1102			

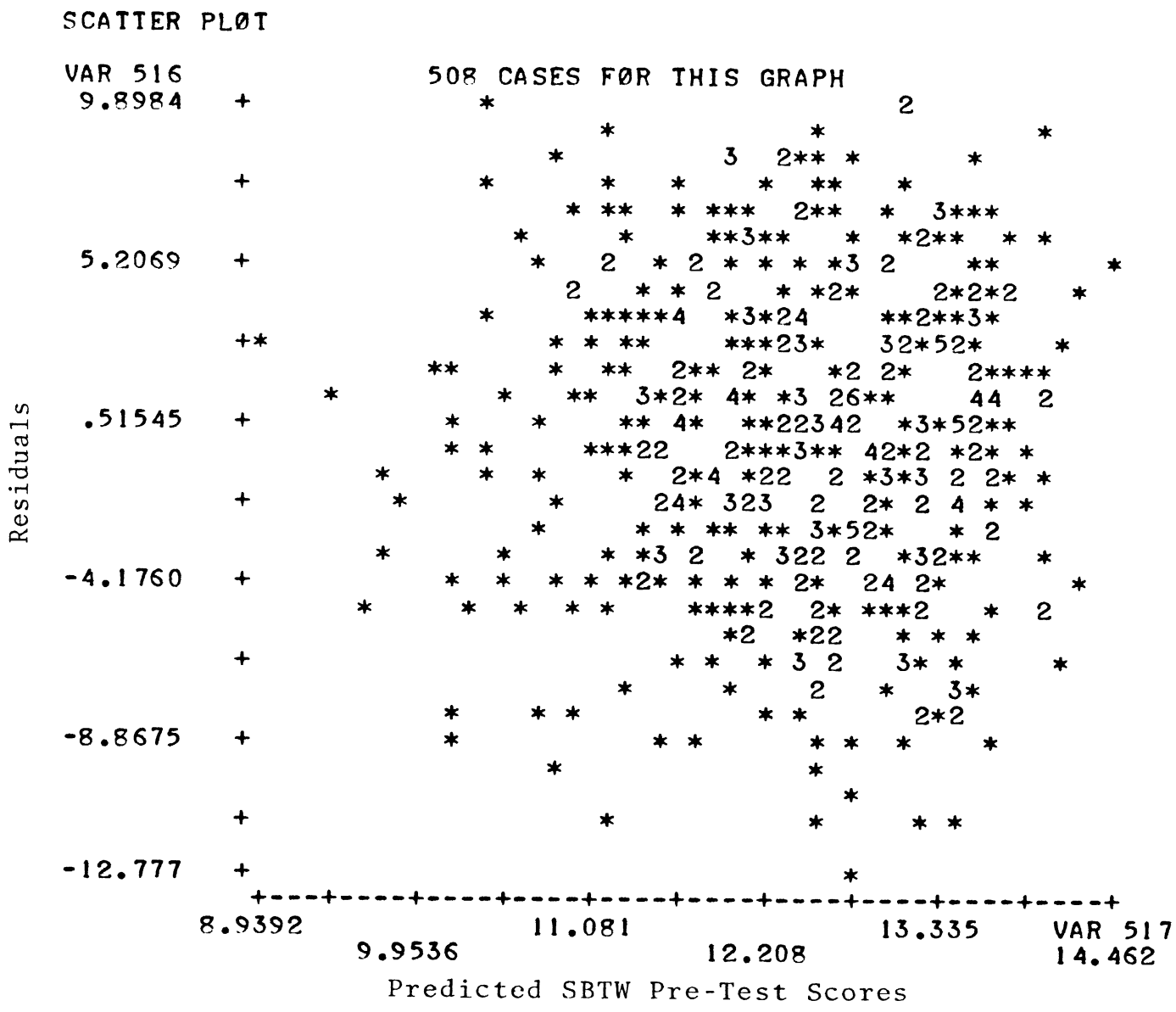
When the residuals of this equation are plotted against predicted values of Y_i (Figure 1), the scatter is essentially random, suggesting that the assumption of homogeneous variances is met.

Given the sample size, and because the R^2 values are so consistently low, further attempts to refine the model, such as polynomial procedures, were not considered to be worthwhile. However, it was decided to check the model using the split-half technique.

3.2 The Split-Half Check Results

The assumption that the MIDAS random number operator (seeded at 3333) is truly random was not tested. Accepting the assumption, the two regression equations generated by the

FIGURE 1: Residuals of Equation 3 Plotted Against Predicted SBTW Pre-Test Scores (Whole Data Set).



random halves appear not to differ on the basis of the techniques used. The results of these two equations are summarized in Table 3.

TABLE 3: Results from Regression Equations Generated from Split-Half Random Samples Using Three Independent Variables.

EQN	N	R^2	S.E.	α
First random sample	260	.037	4.5	.0215
Second random sample	248	.073	4.4	.0003

The two equations are:

$$\text{First sample: } \hat{Y}_i = 12.64 + 0.28082X_1 - 0.06696X_2 + 0.00075X_3$$

$$\text{Second sample: } \hat{Y}_i = 16.936 + 0.06969X_1 - 0.13975X_2 + 0.00178X_3$$

Even though the standard errors are almost identical, the differences in R^2 values warrant some further comparisons.

Plots of residuals against predicted values of Y_i suggest that the assumption of homogeneous variances is equally well met by the two equations, and this also holds true when the equation from the first sample is applied to the second sample.

The Y_i's predicted for the second sample by the second sample equation were plotted against those predicted for the sample sample by the first sample equation. The trend is clearly linear but shows more variability than a similar plot comparing the values predicted on the second sample by the equation derived from the full data set with those predicted by the second sample equation. It will be noted that the two equations in this comparison are closer in form than those in the previous comparison.

It may be concluded, without further analysis, that on

the basis of this rather large data set, is it unreasonable to expect socio-economic status, age and driving experience to predict more than approximately five percent of the variance of scores on the SBTW driving test.

4.0 Summary and Conclusions

4.1 Summary of Findings

The general conclusion that the SBTW score is essentially independent of three variables which are popularly believed to influence this kind of test is very useful in the context of a difficult evaluation project. The assumptions of the regression models (homogenous variances, independence of observations, linear relationship) appear to be met, but the consistently low R^2 values do not warrant further refinement of the models. Similarly, more sophisticated techniques for comparing the split-half regression equations, such as analysis of covariance and the plotting of confidence bands, are unnecessary in this situation. There is considerable value to making explicit the predictive ability of each of the selected independent variables, as part of the validation of instruments of this kind.

4.2 Possibilities for Future Work

The failure of driving experience to have more than a minute effect on SBTW score is somewhat suspect. Some fault may be with the method of collecting that information. An alternative exposure measure might be worth gathering specially if such a test were to be extended to wider applications (in this case it is not so intended).

As a separate study, it would be interesting to follow the same procedure using pre-test/post-test difference scores, although the correlation matrix in Appendix 1 is not promising in this regard.

It would certainly be worthwhile to seek other predictors

of SBTW score. As the larger HSRI study proceeds, it will be possible to look at it in relation to previous and current driving record and the scores of other criterion tests.

REFERENCES

1. Martin E. Lee. "Detecting the Effectiveness of Driver Retraining Among Detroit Car Commuters," Universities Transport Study Group Conference, University College, London, January 3-4, 1973. U.G.T.S. 1973.

2. Philip S. Carroll, Robert E. Scott, Thomas L. McDole, and William L. Carlson. "Acquisition of Information on Exposure and on Non-Fatal Crashes." Volumes I-V, Executive Summary. Highway Safety Research Institute, May 12, 1971. Final Report Under Contract FH-11-7293 DOT.

APPENDIX 8

NON-VOLUNTEER SURVEY
MAILED JULY-OCTOBER 1972

Ford Motor Company

The American Road
Dearborn, Michigan 48121

July 12, 1972

Dear Ford Employe:

Last November, we invited you to take part in the Company's Employe Skilled Driving Program. The offer was extended to a randomly selected sample of 2,000 salaried employes in the greater Detroit metropolitan area.

More than 1,000 volunteers have been assigned to one of four courses now underway. We hope to extend the Program to more employes in the future. Meanwhile, it would be of considerable help to us in further developing appropriate materials if you would answer the two questions on the enclosed form. Your signature is optional.

We would appreciate the return of the form as soon as possible. Please use Company mail.

Thank you for your cooperation.

Sincerely,

Phil Gram, Manager
Employe Skilled Driving Program

Enclosure

MAIL TO: Phil Gram, Manager
Employe Skilled Driving Program
Main Lobby - EAST
World Headquarters

1. Do you feel it is worthwhile for Ford Motor Company to provide skilled driving instruction for its employes?

   Yes _____

   No _____

   Please give your reason for this answer:

   _____

   _____

   _____

   _____

   _____

   _____

2. Why did you decide not to take advantage of this course?

   _____

   _____

   _____

   _____

   _____

   _____

Signature: _____
(Optional)