September 1981
Final Report

DOT HS-806-207

U.S. Department of Transportation
National Highway Traffic Safety Administration

Automobile Driver On-Road Performance Test

Volume I—Final Report

Kenard McPherson
A. James McKnight

National Public Services Research Institute
123 North Pitt Street
Alexandria, Virginia 22314

Contract No. DOT HS-9-02092
Contract Amount $150,043

This document is available to the U.S. public through the National Technical Information Service, Springfield, Virginia 22161

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

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT-HS-806 207		

4. Title and Subtitle
Automobile Driver On-Road Performance Test Volume I: Final Report

5. Report Date
30 September 1981

6. Performing Organization Code

7. Author(s)
A. James McKnight, Kenard McPherson

8. Performing Organization Report No.

9. Performing Organization Name and Address
National Public Services Research Institute
123 North Pitt Street
Alexandria, VA 22314

10. Work Unit No. (TRAIS)

11. Contract or Grant No.
DOT-HS-9-02092

12. Sponsoring Agency Name and Address
National Highway Traffic Safety Administration
400 Seventh Street, S. W.
Washington, D.C. 20590

13. Type of Report and Period Covered
Final Report
3/79 - 9/81

14. Sponsoring Agency Code

15. Supplementary Notes Dr. Stephen V. Versace served as Contract Technical Manager throughout the research and development phase of the project. Mr. Michael F. Smith took over as Contract Technical Manager during the data analysis and reporting phases.

16. Abstract

The Automobile Driver On-Road Performance Test (ADOPT) was developed during a three-phase project. In Phase 1, 51 candidate behaviors were identified and selected with the help of experts in the fields of traffic safety, measurement of driver performance, and driver licensing. In Phase 2, candidate behaviors were subjected to testing to determine their measurability and the best methods for achieving valid and reliable measurement. Relevant criteria for use in assessing performance were established. During Phase 3, the ADOPT was developed and documented in two manuals, one for administrators and one for examiners. The 10-minute test uses an objective scoring system and assesses specific performances at designated locations only. It is administered by one examiner.

The ADOPT was pilot-tested in Oklahoma to assess its reliability and validity. Following the pilot test, the ADOPT was modified to improve overall reliability and validity. The revised version was field tested, again in Oklahoma. Results showed the ADOPT to be a reliable measure of the on-road performance of license applicants. It is primarily a measure of driver skill, which serves as an indirect measure of safe driving practices. Intercorrelation of applicant scores across examiners and across routes exceeds .8 and .7, respectively.

17. Key Words
road test
driver performance
license test
safe driving practices
driving skills

18. Distribution Statement
Document is available to the public throug
the National Technical Information Service
Springfield, Virginia 22151

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	112	

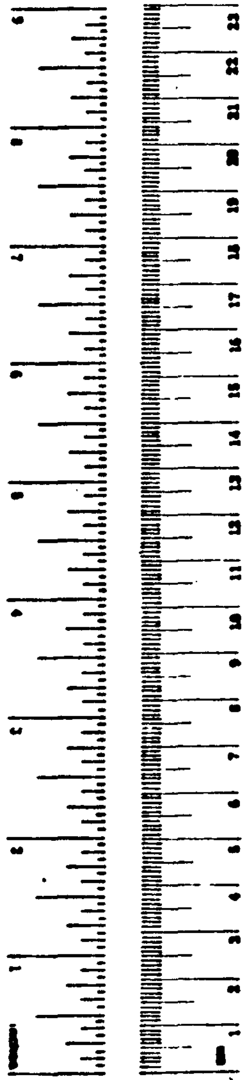
Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

METRIC CONVERSION FACTORS

Approximate Conversions to Metric Measures

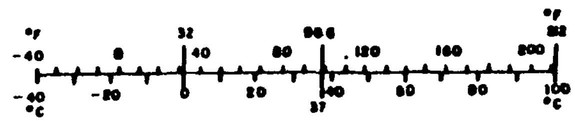
Symbol	When You Know	Multiply by	To Find	Symbol
	LENGTH			
in	inches	*2.5	centimeters	cm
ft	feet	30	centimeters	cm
yd	yards	0.9	meters	m
mi	miles	1.6	kilometers	km
	AREA			
in^2	square inches	6.5	square centimeters	cm^2
ft^2	square feet	0.09	square meters	m^2
yd^2	square yards	0.8	square meters	m^2
mi^2	square miles	2.6	square kilometers	km^2
	acres	0.4	hectares	ha
	MASS (weight)			
oz	ounces	28	grams	g
lb	pounds	0.45	kilograms	kg
	short tons (2000 lb)	0.9	tonnes	t
	VOLUME			
tsp	teaspoons	5	milliliters	ml
Tbsp	tablespoons	15	milliliters	ml
fl oz	fluid ounces	30	milliliters	ml
c	cups	0.24	liters	l
pt	pints	0.47	liters	l
qt	quarts	0.95	liters	l
gal	gallons	3.8	liters	l
ft^3	cubic feet	0.03	cubic meters	m^3
yd^3	cubic yards	0.76	cubic meters	m^3
	TEMPERATURE (exact)			
°F	Fahrenheit temperature	5/9 (after subtracting 32)	Celsius temperature	°C

*1 in ≡ 2.54 (exactly). For other exact conversions and more detailed tables, see NBS Misc. Publ. 286, Units of Weights and Measures, Price $2.25, SD Catalog No. C13.10:286.



Approximate Conversions from Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
	LENGTH			
mm	millimeters	0.04	inches	in
cm	centimeters	0.4	inches	in
m	meters	3.3	feet	ft
m	meters	1.1	yards	yd
km	kilometers	0.6	miles	mi
	AREA			
cm^2	square centimeters	0.16	square inches	in^2
m^2	square meters	1.2	square yards	yd^2
km^2	square kilometers	0.4	square miles	mi^2
ha	hectares (10,000 m^2)	2.5	acres	
	MASS (weight)			
g	grams	0.035	ounces	oz
kg	kilograms	2.2	pounds	lb
t	tonnes (1000 kg)	1.1	short tons	
	VOLUME			
ml	milliliters	0.03	fluid ounces	fl oz
l	liters	2.1	pints	pt
l	liters	1.06	quarts	qt
l	liters	0.26	gallons	gal
m^3	cubic meters	35	cubic feet	ft^3
m^3	cubic meters	1.3	cubic yards	yd^3
	TEMPERATURE (exact)			
°C	Celsius temperature	9/5 (then add 32)	Fahrenheit temperature	°F



PREFACE

This report describes the development and evaluation of an Automobile Driver On-road Performance Test (ADOPT). The work was performed by the National Public Services Research Institute (NPSRI) under contract to the National Highway Traffic Safety Administration (NHTSA), U.S. Department of Transportation (Contract No. DOT-HS-9-02092). Drs. Mark Lee Edwards, A. James McKnight, and Kenard McPherson served as Principal Investigators during the various phases of the project.

The project staff is grateful to Dr. Stephen V. Versace and Michael F. Smith, NHTSA Contract Technical Managers, for their guidance throughout the project. Dr. Versace served as CTM during initial development and pilot testing. Mr. Smith served as CTM during the field evaluation and final developmental work.

We acknowledge the contributions of the Project Advisory Committee members, who graciously provided their time and advice to the project staff. Committee members were: Dr. Mark Lee Edwards, NPSRI (formerly with Texas Transportation Institute); Wayne Green, Nebraska Department of Motor Vehicles; Dr. Margaret H. Jones, University of Southern California; John F. O'Brien, New York State Department of Motor Vehicles, Mike Rudisill, Michigan Department of State; Maj. Howard R. Showe, Maryland Motor Vehicle Administration; and Maj. Thomas Tennery, Oklahoma Department of Public Safety.

We wish to express our appreciation to Rodger Koppa, Texas Transportation Institute, Texas A & M University, who directed subcontract work conducted by TTI in determining the feasibility of and techniques for measuring certain behaviors under consideration for inclusion in the ADOPT.

We also wish to express our thanks to staff of the Oklahoma Department of Public Safety (DPS), who provided support for the pilot and field tests of the ADOPT. In particular, we acknowledge the help of Maj. Thomas Tennery, Capt. Bill Williams, Capt. John Holland, and Lt. Kenneth Thompson, who were in charge of the DPS effort. We are grateful to the examiners from the Oklahoma City and Tulsa/Jenks examination offices, who provided their time and assistance.

Finally, we acknowledge the assistance of the following NPSRI staff members who contributed to the project: Phil Durham, Ruth Freitas, Curtis Goode, Anne Knipper, and Michael Sadof.

The following two volumes complete this report series:

- Automobile On-Road Performance Test (ADOPT), Volume II: Administrator's Manual
- Automobile On-Road Performance Test (ADOPT), Volume III: Examiner's Manual

TABLE OF CONTENTS

Table of Contents (continued)

Table of Contents (continued)

Page

Table of Contents (continued)

Page

INTRODUCTION

All States administer a road test to those persons seeking their first driver's license. Road tests are needed to meet State statutory requirements that applicants demonstrate minimal operating skills. The overwhelming majority of Americans will at some time take road and written driving tests, and it is hard to think of other tests that are as widely accepted as these licensing tests are by those who administer them and those who take them. There are few who question the need to test drivers' knowledge of laws, signs, and safe driving practices or their ability to handle a car in traffic.

What is questioned is the effectiveness of the existing road tests. In the light of increasing costs of administration, some have questioned whether all of the various driving tasks assessed during the road tests really reflect the applicant's ability to drive safely. Several studies have attempted to assess the effectiveness of existing road performance tests by correlating test scores with subsequent accident and violation experience. Campbell (1958), McRae (1968), and Harrington (1973) all found significant but very small correlations. Kaestner (1964), as well as Waller and Goo (1968), found both positive and negative correlations, with results dependent on the age and sex of applicants. Wallace and Crancer (1969) and Dreyer (1976) found no correlation.

In each of those studies, the investigators concluded that the road tests lacked sufficient <u>predictive validity</u> to support their use as a screening device in determining who will be permitted to drive. This was not really surprising because road tests are not aimed at <u>predicting</u> accidents. They are aimed at <u>preventing</u> accidents by assuring that those who drive have the ability to do so safely. The American Association of Motor Vehicle Administrators has stated (1967) that the purpose of the road test is "to assure the applicant's ability to drive safely." That is precisely what existing State road tests have attempted to do. The tests are meant to assure that persons who are licensed have demonstrated their ability is of acceptable quality.

The problem with existing road tests lies principally in the fact that they have not been developed from a systematic analysis of critical driving tasks. They have evolved unscientifically, and their content validity, examiner reliability, and sampling reliability are unknown. What is needed by the States is a road test that is:

- <u>Valid</u>. It must assess those behaviors that are critical to safe operation of an automobile.
- <u>Reliable</u>. Each sample of behavior must provide a reliable estimate of total driving performance, regardless of variation in route or traffic characteristics.
- <u>Objective</u>. The scores applicants receive must depend totally upon performance and must not vary as a function of differences among examiners.

- Feasible. The test must be administrable under constraints imposed by limitations in applicant and examiner time, available manpower, personnel skills, and the characteristics of available routes.

- Safe. The test must not expose the applicant or the examiner to hazards beyond those that prevail in everyday driving.

- Effective. Administration of the test should result in improved safety, evidenced by a reduction of accidents.

This report describes a study that was conducted to develop such a test: the Automobile Driver On-road Performance Test--ADOPT.

PROJECT PHASES

Project work was essentially divided into three phases.

Phase 1--Identification and Selection of Candidate Behaviors

During this initial phase, project staff conducted a review of relevant literature and examined the research performed in the development of other on-road performance tests. A preliminary list of candidate behaviors for the ADOPT was identified. Following review of the list by groups of experts, a panel met to discuss the list and make a final selection of candidate behaviors.

Phase 2--Study of Measurement Methods

In this phase, certain candidate behaviors were subjected to testing in order to determine their measurability, as well as the best methods for achieving valid and reliable measurement. The study also established relevant criteria for use in assessing performance.

Phase 3--Development and Testing of the ADOPT

In this last phase, a preliminary ADOPT was prepared and pilot tested to assess its validity and reliability. Following revisions based on results of the pilot test, the ADOPT was field tested and developed in final form. The test was documented in an Administrator's Manual[1] and an Examiner's Manual[2].

The following sections of this report describe the work carried out during the three phases of the project.

---

1 Automobile Driver On-Road Performance Test (ADOPT), Volume II: Administrator's Manual.

2 Automobile Driver On-Road Performance Test (ADOPT), Volume III: Examiner'S Manual.

IDENTIFICATION AND SELECTION OF CANDIDATE BEHAVIORS

During this initial phase of the project, project staff identified candidate behaviors for inclusion in the ADOPT. They also reviewed the research conducted in the development of other road tests and reviewed all relevant literature. Once a list of candidate behaviors had been developed, groups of experts screened the list. Following this, a panel was convened to make the final selection of behaviors.

IDENTIFICATION OF CANDIDATE BEHAVIORS

To identify candidate behaviors for an Automobile Driver On-road Performance Test (ADOPT), project staff reviewed the Driver Education Task Analysis, Volume 1, Task Descriptions (McKnight and Adams, 1970). This report contains detailed descriptions of approximately 1,700 behaviors required of passenger car drivers. The analysis separates the tasks into on-road and off-road behaviors as follows:

On-Road Behaviors

- Basic Control Tasks
- General Driving Tasks
- Tasks Related to Traffic Conditions
- Tasks Related to Roadway Characteristics
- Tasks Related to the Environment
- Tasks Related to the Car

Off-Road Behaviors

- Pretrip Tasks
- Maintenance Tasks
- Legal Responsibilities

Initial Screening of Behaviors

Many of the behaviors listed were clearly unsuitable for inclusion in an on-road performance test. To identify behaviors that might be suitable, project staff screened the 1,700 tasks to eliminate those that were:

- Nonoperating--Behaviors that were not directly related to the operation of the automobile, e.g., maintenance, pushing and towing, loading.
- Infrequent--Behaviors required for response to situations that rarely occur, e.g., avoiding a collision.
- Unsafe--Behaviors that would create undue risks for the applicant and the examiner or that could result in damage to the applicant's vehicle, e.g., off-road recovery.
- Safety-Unrelated--Behaviors that were not related to safety, e.g., trip planning.

Eliminating behaviors in those four categories resulted in a preliminary list of 165 candidate behaviors for further consideration. The complete list is provided in Appendix A.

Classification of Behaviors

The behaviors surviving the preliminary screening were separated into the following five categories.

1. Driver/Vehicle Readiness--Behaviors required prior to driving.

2. Vehicle Control--Behaviors involved in simple control of vehicle motion.

3. Vehicle Maneuvering--Behaviors involved in control of vehicle motion during specific maneuvers.

4. Interacting with the Highway/Traffic Environment--Behaviors required for effective interaction with components of the highway/traffic environment.

5. Interacting with Highway/Traffic Hazards--Behaviors of a high skill level that are involved in interacting with highway/traffic hazards.

The five categories were used primarily to facilitate further development work. The categories permitted the behaviors to be grouped logically and provided an efficient means of locating individual behaviors, since without some type of division a search of the entire list would have been required in order to locate a specific behavior.

The categories can be further grouped into "Skills" and "Practices" as follows:

Skills

Vehicle Control
Vehicle Maneuvering
Interacting with Highway/Traffic Hazards

Practices

Driver/Vehicle Readiness
Interacting with the Highway/Traffic Environment

Behaviors in the "Skills" category are those in which successful performance is dependent primarily upon the mastery of manipulative and perceptual skills. Those in the "Practices" category are not in and of themselves heavily dependent upon skills (anyone can signal a turn). Rather they reflect day-to-day ways of behaving. The two categories of behavior were believed to differ substantially in their implications for measurement procedures. These implications will be discussed more fully later in the report.

REVIEW OF SELECTED DRIVER PERFORMANCE TESTS

A review of selected tests of driver performance was undertaken to identify:

- Operational factors that should be taken into consideration in developing a licensing test of driver performance.

- Candidate behaviors and measurement techniques appropriate for a test of driver performance within the context of driver licensing.

The review was limited to performance tests meeting the following criteria:

Objectivity of Measurement--Driver performance is objectively measured and scored. Tests utilizing performance measures based solely on the examiner's subjective interpretation of the driver's performance were excluded from the review.

Demonstrated Reliability--Only tests with documented reliability coefficients for such test characteristics as interexaminer agreement, sampling reliability, and route-to-route reliability were reviewed.

Clearly Defined Test Situations--Those tests with inadequate (or no) descriptions of test situations, scoring methods, or criteria were excluded from the review.

Three tests of driver performance were selected for review on the basis of these criteria. They were:

University of Southern California Safe Performance Test

Michigan State University Driver Performance Measure

Motorcyclist In-Traffic Test

A general description of each test, outlining test content, administration requirements, and measures of reliability, is provided below.

Safe Performance Test (SPT)

This test of driver performance was developed by the University of Southern California, under NHTSA sponsorship, as an intermediate criterion for evaluating the Safe Performance Curriculum for driver education students. It has been designed to measure selected behaviors considered critical to the safe operation of a motor vehicle, as defined in the Driver Education Task Analysis (McKnight and Adams, 1970). Only those behaviors rated as either "high" or "moderate" in criticality are included.

Test Content

A total of 30 individual test items, grouped into several major categories of driver performance, comprises this test. The more important of these major categories are:

> Vehicle Path--Position of the vehicle in travel lanes during left and right turns, and lane changes.
>
> Speed--Vehicle speed during turns, lane changes, and straight roadway sections.
>
> Observation--The driver's use of head/eye checks to observe traffic conditions during turning, intersection approaches, and lane changes.
>
> Mirrors--Use of mirrors prior to lane changes and turns.
>
> Gap Judgment--Judgment (and maintenance) of proper gaps when following other vehicles and traversing intersections.

Additional test items provide for measures of more basic driver skills involved in: backing, turnarounds, stopping, preoperation, and shutdown of the vehicle. Separate measures of the driver's response to hazards that may occur during the test, as well as the frequency with which the examiner must take control of the vehicle, are also provided.

Test Administration

The total time required to administer this test is approximately 30 minutes. Two examiners are needed. A front seat examiner is responsible for scoring the driver's response to hazards, maintaining control of the vehicle, and administering route instructions. The second examiner, located in the rear seat, scores the driver's performance in all other test situations. Approximately 40 hours of training are required to administer this test.

Scoring

The driver's response to each test item is scored as either correct, incorrect, or unobservable. Individual scores are recorded by coding a schematic of the test route. Item scores are summed to provide "subtest" scores, which are then summed to provide an overall test score.

All item scores are calculated in terms of "percentage of correct responses" to correct for differences in the frequency with which individual test items are encountered in the test route.

Test Reliability

Examiner agreement, one measure of interexaminer reliability, is approximately 80% for total test score. The percent agreement for subtest scores ranges from 76% to 82%.

Sampling reliability, measured by correlating examiner scores for the same applicant on the same route over two administrations of the test, is:

- Total Test Score: r = .80
- Subtest Scores: r = .51 to .90
- Test Items: r = .10 to .85

No measures of route-to-route reliability have been reported for this test.

Driver Performance Measure (DPM)

This test was recently implemented by the Michigan Department of State, Driver License Division, as a replacement of their road test for initial applicants. It is based on a concept for observing and measuring driver behavior originally developed by Forbes (1975). A principal assumption underlying this test is that valid measures of driver performance must be sensitive to the dynamics of traffic and roadway conditions and the driver's reaction to them.

Little has been written regarding specific behaviors or measurement techniques employed in this measurement approach. Individual test situations are couched in terms of BETSS (behavioral-environmental-traffic-situational sequences). Individual BETSS are comprised by a number of sub-BETSS (approximately three per BETSS). The examiner, when scoring driver performance in response to these BETSS, attempts to gauge the driver's overall response to the various characteristics of each BETSS and sub-BETSS.

Test Content

The content of individual BETSS and sub-BETSS varies. The overall suitability of the driver's response to these various test situations is assessed in consideration of the following factors:

- Driver search behavior
- Speed control
- Direction control
- Pattern of driver's response
- Total score

The examiner scores driver performance as either "suitable" or "unsuitable" given the overall nature of the driver's response to the individual conditions that arise within a given BETSS.

Test Administration

Approximately 15 to 30 minutes are required to administer the DPM. A single examiner is required, located in the passenger side of the front seat. Scoring is accomplished by checking the appropriate performance indicator on a schematic of the test route.

Scoring

The driver's overall response pattern is scored as either "suitable" or "unsuitable." Additionally, three general observations are made throughout the test route between individual BETSS. These are:

- near-miss accidents.
- hazardous moving violations.
- performance below minimum acceptable levels.

Performance is evaluated at three levels: item score values, subscore values, and total score. Scores are derived by simply totaling the number of suitable responses. Approximately 120 hours of training are required for administering this test.

Test Reliability

Interexaminer reliabilities for item scores, subscores, and total scores were determined by utilizing two examiners to administer the DPM to the same driver over a single route. The reliability coefficients obtained for each scoring element are:

- Individual score values: r = .56
- Subscore values: r = .58
- Total score: r = .59

No estimates of sampling reliability have been determined as yet for this particular test of driver performance.

Motorcyclist In-Traffic Test (MIT)

The MIT was developed by tne National Public Services Research Institute under NHTSA sponsorship. The objective of this particular test is to provide an in-traffic measure of motorcycle operator performance as an alternative to an off-street test previously developed. Its intended use is as a licensing test only, for initial or renewal applicants.

Test Content

The MIT provides for testing of applicant performance in 21 different highway and traffic situations. Both basic skills and safe operating practices are included. The individual behaviors comprising this test are grouped in five major categories:

Observation--The rider's application of head/eye checks to detect the position and actions of other drivers.

Signaling--The application of turn signals to communicate the rider's intentions to other drivers.

Longitudinal Positioning--Positioning of the motorcycle in response to leading vehicles and following vehicles.

Lateral Positioning--Positioning of the motorcycle within lane boundaries during turns and over-the-road riding situations.

Gap Selection--Judging intervehicle gaps at intersections.

Each of these major categories of driver behavior subsumes a number of individual behaviors. For example, Observing behaviors are subdivided into:

- Observing ahead.
- Observing to the sides.
- Observing signs, signals, and markings.
- Observing travel restrictions.

Basic skill measures included in the test are:

- Starting the motorcycle.
- Putting the motorcycle in motion.
- Maintaining directional control.
- Stopping the motorcycle.

The test is structured so that the applicant's performance is scored in the light of responses to specific behaviors imbedded in selected traffic maneuvers. No overall assessment of the driver's performance of the maneuver is made by the examiner. Instead, maneuvers are selected to maximize the likelihood that specific behaviors will be exhibited by the driver. Only these behaviors are scored, regardless of what other behaviors might be elicited during the driver's performance of the maneuver.

Test Administration

Approximately 10 to 15 minutes are required to complete the actual on-road testing activity. A single examiner is needed. The applicant is followed by the examiner in a separate vehicle. Route instructions are communicated to the applicant via a one-way FM radio. The examiner drives the following vehicle, administers instructions, and scores applicant performance. Where required, "safe" zones are established in the route to allow the examiner to pull off the roadway to complete performance checks.

The applicant's performance is scored for each behavior as follows:

Yes--Applicant performed to specified criterion level.

No--Applicant did not perform according to criterion.

N/A--Opportunity to observe behavior was not available.

The applicant's total score is derived by totaling the number of correct responses. Approximately 16 hours of training are required to administer this test.

Test Reliability

Interexaminer reliability, determined by correlating examiner judgments for the same rider over a single route, ranges from an r of .32 to .87 for major performance categories. Total score interexaminer reliability is approximately .60.

Sampling reliability (obtained by correlating a single examiner's assessment of performance over two routes for the same applicant) is approximately r = .62. The correlation of the MIT total score with scores on an off-street skill test--the Motorcycle Operator Skill Test (MOST)--is approximately .50.

Summary

The most important characteristics of each of the performance tests just described are summarized in Table 1 on the following page.

The various methods for assessing driver performance identified in this review of selected driver performance tests were examined in relation to the candidate behaviors for the ADOPT. This helped in the determination of:

- Alternative measurement approaches--for individual behaviors to be included in an on-road test of driver performance.
- Measurement techniques--to be developed or refined in those instances where no existing measure was available.
- Behavioral Definitions--required to objectively define performance criteria for specific driver behaviors.

TABLE 1

SELECTED PERFORMANCE TEST CHARACTERISTICS

TEST	BEHAVIORAL CATEGORIES	NO. OF EXAMINERS	EXAMINER TRAINING HRS.	TEST DURATION	PERFORMANCE SCORING	TOTAL SCORE	EXAMINER AGREEMENT	SAMPLING RELIABILITY
SPT	Vehicle path Speed Observation Mirrors Gap judgment Car handling skills	2	40	30 min	Maneuvers scored: "Correct" "Incorrect" "Unobserved"	Percent correct manuevers	80%	.80
DPM	Search behavior Speed control Direction control Pattern of responses	1	120	15-30 min	Behaviors scored: "Suitable" "Unsuitable"	Number of suitable behaviors	r = .59	Not available
MIT	Observation Signaling Longitudinal positioning Lateral positioning Gap selection Basic handling skills	1	16	10-15 min	Behaviors scored: "Yes" "No" "N/A" against specific criteria	Percent "Yes"	.60	.52

REVIEW OF HUMAN PERFORMANCE ASSESSMENT LITERATURE

A review of literature relevant to the assessment of human performance (particularly driver performance) was undertaken as a preliminary step in the development of the ADOPT to identify individual behaviors and techniques for measuring behaviors appropriate for application in an on-road test of driver performance. Particular emphasis was placed on those studies addressing:

> Psychological Characteristics of Drivers--The relationship between personality characteristics of the driver and accident involvement.
>
> Psychomotor Abilities--The driver's capacity to control speed, path, and direction of a motor vehicle.
>
> Perception and Vision--Visual acuity and the use of proper search patterns.
>
> Safe Operating Practices--The utilization of accepted principles of safe motor vehicle operation in the driving task.

The review was confined to studies attempting to quantify, either by direct measurement or observation, specific driver behaviors. Publications of a purely descriptive nature and those presenting nothing more than conceptualizations of driver performance characteristics were excluded from this review.

Psychological Characteristics

A number of studies have investigated the relationship between accident involvement and basic personality or attitudinal characteristics of the driver, most being devoted to simply describing the personality characteristics and the nature of their relationship. Few studies have attempted to determine, in an experimental setting, the causal relationships between specific personality characteristics and accident involvement.

Studies such as those performed by Tillman and Hobbs (1949), Thorndike (1951), Selzer and Payne (1962), Tabachnick et al. (1966), Brown and Bonhert (1968), Selzer, Rogers, and Kern (1968), Crancer and Quiring (1970), McMurray (1970), Haviland and Wiseman (1974), and Phillips (1977) are representative of the former approach to studying the relationship between personality characteristics and accidents. In these studies, and numerous others, significant correlations have been found between such driver characteristics as criminal record, credit rating, employment history, personal problems, suicidal tendencies, etc. While studies of this type provide some insights into the interaction between personality characteristics and driving styles, they are purely descriptive in nature and provide no direct evidence of any "causal" relationship between personality attributes and accident risk.

Quenault (1968) attempted to assess in a more direct fashion the influence of personality characteristics on specific driving styles, and was able to identify two basic personality types associated with high accident involvement:

(1) Disassociative "active"
(2) Disassociative "passive"

Drivers classified as disassociative active were characterized by their general edginess and impatience while driving. Quenault found their behavior to be unpredictable, such drivers often responding differently in similar situations or similarly in vastly different driving situations. Disassociative passive drivers, on the other hand, were found to behave in the same general manner regardless of the particular driving situation being faced.

Edwards (1972) correlated responses to a pencil-and-paper personality inventory with the accident history of approximately 1,000 drivers under the age of 22. The results of a factorial analysis of data indicated that accident-involved drivers did possess some basic personality characteristics that distinguished them from their "accident-free" cohorts. However, subsequent analyses concluded these traits were as descriptive of the "youth" population in general as they were of accident-involved drivers, indicating a possible confounding of the relationships between personality characteristics and accident involvement as a function of age.

Perhaps the most thorough investigation of the role personality factors play in accident involvement was undertaken by Mayer and Treet in 1977. The initial study focused on two groups of college students (accident-involved and accident-free) matched for age, sex, and annual mileage. All accident group members had three or more accidents during the prior three-year time period. A battery of tests measuring some 20 personality characteristics was adminstered to members of both groups. Six tests were identified as capable of distinguishing the control group from the accident group. A follow-up study using a cross-validated prediction model permitted the authors to correctly identify 12 of 14 other drivers as being either accident-involved or accident-free.

These results are supportive of those obtained in other studies; namely Haviland and Weismann (1974), Schmidt et al. (1976), and McGuire (1976). All concluded that reliable discriminations between accident-free and accident-involved drivers could be made on the basis of personality characteristics measurable with pencil-and-paper instruments.

Personality Characteristics and Licensing

The difficulties associated with utilizing personality characteristics as a means of determining driver "fitness" within the context of driver licensing is obvious. State-operated driver licensing programs are designed (and required by law) to assess the skills and abilities of drivers to operate a motor vehicle safely. While personality characteristics are related to accident involvement, in and of themselves they do not comprise a measure of performance ability. Furthermore, the evidence amassed to date has failed to establish any causal relationships between personality characteristics and driving ability. In fact, some evidence suggests those relationsnips that have been found to exist merely reflect the obvious, i.e., personality characteristics influence virtually all aspects of an individual's life style. As such, personality characteristics related to accident involvement are in turn related to the probability of divorce, the probability of engaging in antisocial behavior, etc.

Psychomotor Abilities

The driver's ability to control vehicle path, speed, and direction and other basic vehicle control abilities have been assessed in a variety of experimental settings. For the most part, studies of the psychomotor abilities of drivers have attempted to develop objective descriptors of basic vehicle control behaviors. These studies investigated such aspects of vehicle control as: steering behavior, speed selection and maintenance, lateral positioning, and headway estimation.

In a majority of these studies, efforts have focused on the measure of one (or only a very few) of these basic vehicle control characteristics.

Comparisons of Novice and Experienced Drivers

Greenshields (1963) authored one of the initial attempts to quantify basic vehicle control behaviors. In this particular study, the frequency and magnitude of steering wheel reversal among novice and experienced drivers in a normal over-the-road driving situation was examined. Novice drivers were found to require more frequent steering inputs to maintain proper road position when compared to experienced drivers.

A later study undertaken by Greenshields and Platt (1964) compared the frequency of steering reversals, brake pedal applications, and accelerator use among experienced and novice drivers. Significant differences were found among these two driver populations for each of these measures. In all cases, novice drivers were found to exhibit more frequent control inputs (reversals) than experienced drivers. These basic control abilities were subsequently found to be related to accident invovement; better control capability (as measured by these variables) positively correlated with lower accident rates.

Kimball, Elingstad, and Hagen (1970) assessed the basic vehicle control abilities of novice and experienced drivers in a driving simulator. Only two categories of basic vehicle control abilities were assessed: speed control and steering control. Experienced drivers were found to exhibit significantly superior performance as evidenced by fewer steering inputs to maintain vehicle path, lower tracking error rates, fewer speed changes, and little variability in lateral positioning.

A study by Quenault and Parker (1973) identified two variables (average speed and incidence of poor vehicle control) to be significantly related with age of the driver, and thus driving experience. Speed was found to increase with experience, while incidences of poor vehicle control decreased. The authors concluded that mastery of basic vehicle control skills, as measured by the frequency of vehicle control errors, permitted the experienced driver to operate at higher speeds with lower error rates.

These results have been confirmed in other studies (Edwards et al., 1966; Zell, 1969; Harootyan, 1969; Shinar, McDonald and Treat, 1978; and others. Collectively, they demonstrate that observable and substantial differences exist in the basic vehicle control abilities of novice and experienced drivers.

Comparison of Accident-Free and Accident-Involved Drivers

Crancer (1968) employed a driving simulator to measure five types of driver basic control errors: speeding, steering, braking, accelerating, and signaling. Resulting measures of these errors were found to be positively correlated with accident involvement, providing additional evidence in support of research conducted by Greenshields, Platt, and others.

The results of in-depth analyses of selected traffic accidents as reported by Shinar, McDonald, and Treat (1978) provide further evidence of the relationship between basic skill abilities and accident involvement. In comparing accident causes observed for experienced and inexperienced drivers, it was found that inexperienced drivers were more likely to be involved in accidents contributable to improper directional control of the vehicle.

Acquisition of Basic Vehicle Control Skills

Studies of the rate at which basic vehicle control skills are acquired indicate that mastery is attained at relatively low levels of driving experience. This is best exhibited in a study conducted by Mourant and Rockwell (1970). Novice drivers were tested before, during, and after driver training. A majority of drivers were found to have mastered a majority of the basic vehicle control skills measured. Examples include speed control, headway estimation, and car following. Most were mastered within a two-to-three-week period following the inception of training. It was noted, however, that those driving tasks requiring substantial decision-making skills required longer time periods for acquisition. In fact, several had not been mastered at the completion of the study.

Basic Vehicle Control Skills and Licensing

The assessment of basic vehicle control skills within the context of a road test of driver performance would appear valid from at least two perspectives.

(1) Available evidence suggests that a mastery of basic vehicle control skills permits the driver to devote "increased attention" to decision-related driving activities. Examples include adjusting speed to traffic conditions, observing traffic, maintaining adequate intervehicle separation, and navigating.

(2) Inadequacies in basic control capability may or may not precipitate accident situations. No direct evidence is available. It is, however, these skills that permit the driver to ultimately avoid or minimize the consequences of an accident once an accident-producing situation arises.

Perception and Vision

The relationship between perception/vision and accident involvement has yet to be well established. This is due not so much to the technical difficulties encountered in measuring the perceptual and visual characteristics of drivers, but the inability to vary such factors in an experimental setting. Virtually all studies in this area of driver performance have relied on correlational techniques to investigate relationships between these factors and accidents.

Visual Acuity

Evidence of the role visual acuity plays in accident involvement has been provided in the studies conducted by Silver (1936), Lauer (1937), Lauer et al. (1939), Cobb (1939), Fletcher (1949), Henderson et al. (1971), and others. These studies have identified small but significant correlations between various measures of visual acuity (both static and dynamic) and accident involvement. Henderson, Burg, and Brazelton (1971) measured 200 drivers with at least 20/40 vision as measured by Snellen acuity, hexobar acuity, and checkerboard acuity. Snellen acuity was found to be significantly correlated with accidents, though the correlation was small ($r = .15$). These results were essentially confirmed in a later study by Henderson and Burg (1974), although significant correlations between Snellen acuity and accident rates were found only for drivers 25 to 49 years of age.

Hofstetter (1976) analyzed the role of binocular visual acuity in accidents. His sample consisted of approximately 14,000 drivers across 27 States. The percentage of drivers with poor visual acuity who had three or more accidents was approximately two times greater than the percentage of drivers with good visual acuity for all age groups in this sample.

Visual Search

The requirement for sophisticated equipment and techniques to accurately measure driver visual search patterns has limited the investigation of this behavioral characteristic. Studies by Zell (1969) and Mourant and Rockwell (1970) utilizing eye movement recording equipment have documented significant and unique differences in the search and scan patterns of novice and experienced drivers. In general, these results indicate that novice drivers:

- Rely less on peripheral vision.
- Scan for shorter distances down the roadway.

- o Scan unsystematically.
- o Fixate on nonrelevant cues for potential hazards.

Implications for Road Testing

The results of virtually all studies of driver search and scan behavior seem to be of value in the development of a road test of driver performance. The principal difficulty in applying much of this research to the testing of drivers is the requirement to employ sophisticated measurement instruments and techniques to measure such characteristics of visual search as:

- o Fixation points
- o Scan patterns
- o Fixation time

All evidence suggests that visual search behavior is of sufficient importance to warrant the development of less sophisticated techniques (relying solely on unaided observation) to assess driver visual search behavior. While many techniques for doing so have been developed and remain in use (principally in driver education), little evidence as to their objectivity and reliability exists.

Safe Operating Practices

Studies devoted to the use of safe operating practices represent some of the first attempts to quantify driver performance characteristics. The studies by Quenault and Parker (1973), Greenshields and Platt (1964), and Greenshields (1963) cited earlier also attempted to develop techniques for defining and subsequently measuring safe operating practices.

In one of the early studies concerned solely with safe operating practices, conducted by Edwards and Hahn (1964), drivers were filmed from a following vehicle. Behavioral errors were recorded by a group of law enforcement officers. Common errors in safe operating practices identified included failure to signal, failure to maintain lane position, and failure to stop completely. Unfortunately, an analysis of the relationship between these errors and previous accident and/or coniction involvement produced no significant correlations. A replication of this study (Edwards and Hahn, 1970) using a different technique for assessing safe operating practices produced similar results.

More recently, Lohman et al. (1976) observed "unsafe driving actions" at selected high accident locations. The frequency with which these unsafe driving actions were observed was combined with tne frequency of accidents at these locations to derive an index of relative risk. Among those behaviors having the highest risk indices were: turning in front of oncoming traffic, pulling out in front of oncoming traffic, following too closely, ignoring traffic controls, driving left of center, and speeding.

A study of driver signaling behavior at intersections by Barch (1958) found that drivers signaled right turns about half the time and left turns about two-thirds of the time. No analysis of the relationship between these errors in safe operating practices and accidents was performed.

Implications for Licensing

The results of these studies provide evidence to support the inclusion of measures of safe operating practices in an on-road test of driver performance. Indeed, the majority of road tests developed to date focus primarily upon this aspect of driver performance.

The principal difficulty involved in including measures of safe operating practices in any test is that of specifying the behavior precisely in the light of the dynamic nature of the driving task. The requirement to engage in and properly execute safe operating practices varies as a function of the situation to which the driver is responding at the moment. The transient nature of the situations makes difficult the quantification of safe operating procedures in a manner that permits their reliable and accurate measurement.

Conclusions

The results of this literature review provide evidence to support the development of specific measurement techniques applicable in an on-road test of driver performance. To a great extent, however, the results of these studies (and their application to the development of an on-road test) are limited due to their measurement requirements, low validity, and lack of acceptability to license agencies.

Mesurement Requirements

Many of the behaviors identified as appropriate for inclusion in an on-road test of driver performance require the use of sophisticated measurement instruments and techniques which cannot easily be used in an applicant's motor vehicle. Furthermore, the interpretation of these measures requires either considerable time or sophisticated analysis techniques, neither of which is appropriate for an on-road test of driver performance within the context of driver licensing.

Validity

The correlations between most measures and accidents are too small to justify use of the measures on the basis of their predictive validity. A need for high predictive validity is not required under two conditions:

> Content Validity--If the measure is one that is generally accepted as defining safe vehicle operation--such as signaling, checking mirrors, and lanekeeping--it has content validity and evidence of an association with accidents is not demanded.
>
> Modifiability--Measures of variables that are amenable to change are generally acceptable even in the absence of a strong relationship with accidents if applicants can overcome the problem.

Unfortunately, many of the measures used in previous research do not fulfill these conditions. Certain of the psychophysical variables studied lack the apparent relationship to safety needed to be accepted as content valid. One of the reasons that personality or personal history measures are not used is that, despite their correlation with accidents, they are not amenable to change. The applicant cannot be told to "go home and practice and then come back again."

Acceptability

Certain measures of driver performance are simply not acceptable for use in a road test that is given as a part of the licensing process. In order to use the test as the basis of determining eligibility for a license, the performances in the test must have the following characteristics:

> Causal Relationship--Each performance must be causally related to the safety of the driving public. Performances cannot be included because they are correlated with accidents if a causal relationship cannot be assumed. For example, while steering reversals may be correlated with accidents, they do not in any way contribute to them.
>
> Legal Basis--The law in most States requires that license tests limit measures of performance to those that are required under the law. This would eliminate measures of personality, regardless of any causal relationships that they may have with accidents.
>
> Time--The time required to employ some road testing procedures greatly exceeds that available to license stations. For routine road testing of initial applicants few license agencies will accept a test that requires more than 10 minutes to administer.

EXPERT REVIEW AND RATING

The next step in the development of the preliminary ADOPT was a review by selected experts of the 165 candidate behaviors. The purpose of this review was to eliminate additional behaviors so as to arrive at a manageable number of appropriate behaviors from which to construct the on-road test. Three groups of five expert reviewers assessed each behavior. One group rated the behaviors for criticality to safety, another for testability, and the third for acceptability. Members of the three groups are listed in Appendix B.

Criticality to Safety

Traffic safety professionals (experts in behavior research, accident investigation, and driver education) reviewed the list for criticality to safety. They rated each of the behaviors on the basis of its importance in the prevention of accidents, i.e., its relationship to highway safety.

Factors influencing criticality are:

Behavior Frequency--the frequency with which automobile operators are called upon to exhibit a behavior.

Error Probability--the likelihood that a behavior will be performed incorrectly when it is required.

Accident Likelihood--the likelihood that an accident will occur if a behavior is incorrectly performed.

Accident Severity--the extent of loss (in terms of property damage, injury) likely to be sustained in an accident resulting from failure to perform a behavior correctly.

Testability

Testability was assessed by experts in measurement of driver performance and development of driver performance tests. These reviewers rated each behavior on the basis of the degree to which it can be accurately tested in an operational environment (during an on-road test).

Factors influencing testability include:

Reliability--the likelihood that situations calling for a behavior will arise frequently enough to provide reliable measurement.

Observability--the extent to which a behavior is capable of being measured by observation, either directly or indirectly through observation of associated behaviors.

Objectivity--the degree to which a behavior can be measured objectively by examiners.

Acceptability

Driver licensing professionals comprised the group of reviwers who assessed the list of behaviors for acceptability. They rated each behavior in terms of its acceptability to licensing agencies in the light of known constraints.

Considerations were:

Safety--the degree to which a behavior can be tested without risk of injury or property damage.

Legality--the degree to which a behavior and the testing of a behavior are permissible within the laws and regulations governing licensing and the operation of motor vehicles.

Policy, Personnel, Funds--the degree to which the testing of a behavior is acceptable in the light of existing policy and available personnel and funds.

Public Attitude--the degree to which the public will accept the inclusion of a behavior in a road test.

Review and Rating Procedures

A modified Delphi technique was used for the review and rating in an attempt to achieve consensus among the members of each group of reviewers. Reviewers used a five-point rating scale to assess each candidate behavior, "1" being the highest rating. For example, behaviors of the highest criticality were to be rated "1," and those of very low criticality were to be rated "5."

Three rating rounds were conducted by mail. Members of each of the three groups of reviewers were provided with the list of 165 candidate behaviors; criteria for determining the degree to which a behavior was critical, testable, or acceptable; and an explanation of the five-point rating scale. Following their completion of the first rating round, reviewers returned their lists to NSPRI, where they were examined for intragroup agreement.

Behaviors on which there was substantial disagreement were noted, and in a second rating round, reviewers were asked to reconsider their assessment of those behaviors. At this time, reviewers were provided with the round-one ratings assigned to those behaviors by other members of their review group. Reviewers were asked to provide justification for their assessments where they differed substantially from those of other reviewers.

Results of the second rating round and the justifications provided by reviewers were analyzed by staff. In the third and final rating round, reviewers were again asked to reconsider their assessment of those behaviors on which there remained a lack of consensus and at this time were provided with summaries of the justifications given by other members of their group. The results of the third round were then averaged by project staff to obtain scores for each behavior in the three areas of criticality, testability, and acceptability. These scores were also multiplied together to create a "total" score for each behavior.

Results of the Rating Process

The Criticality Review Group achieved a fairly high degree of agreement on all behaviors. In the areas of Testability and Acceptability, however, it was impossible to achieve a sufficiently high degree of agreement, even after three rounds of rating, to aid in determining which behaviors would be eliminated. In spite of the lack of consensus, the results of the rating process were helpful because they clearly identified the critical issues and problem areas.

Testability

In the case of testability, lack of consensus was the result of differing interpretations on the part of reviewers relating to real-world and test performance. For example, a behavior in the category of Driver/Vehicle Readiness was "has not taken drugs or alcohol." Certain reviewers said that this was highly testable because it is easily determined whether an applicant is under the influence of alcohol. Others said it was not testable because real-world behavior could not be assessed; i.e., it could not be determined whether the applicant would consume alcohol prior to real-world driving. This problem of interpretation meant that some reviewers rated the list for testability at the time of the examination, while others considered testability in relation to real-world driving. This affected nearly all of the behaviors in the Driver/Vehicle Readiness category.

Acceptability

In the area of acceptability, the dividing issue was that of legality. Some reviewers felt that there was no point in testing for something that was not required by law because a license cannot be withheld from an applicant on the basis of his inability or failure to do something that is not required by law. Others felt that anything related to safety was acceptable. As an example, those who held the former opinion rated "fastens safety belts" as of low acceptability. Those who held the latter opinion felt that the behavior was highly acceptable.

SELECTION OF BEHAVIORS

Selection Panel

A panel was convened to resolve fundamental issues identified by the rating process and to make a final selection of behaviors for the preliminary ADOPT. The panel consisted of project staff, NHTSA representatives, and expert consultants. The names of selection panel members are provided in Appendix C.

The selection panel completed their work during a two-day meeting. Panel discussion was based upon:

- o the ratings of the 165 behaviors by the three groups of reviewers.
- o information on the behaviors comprising the on-road performance tests referenced earlier in this report and on the methods of measurement used to assess those behaviors.
- o background information on earlier project work and other resource information provided by project staff.

Following their deliberations, the panel selected from the list of 165 behaviors those it considered feasible for inclusion in an on-road test. Behaviors not considered feasible were eliminated as noncritical, nontestable, or nonacceptable. Some behaviors could have been eliminated for more than one reason, but in such cases, only the principal reason for elimination was indicated. The objective of the panel's work was the selection of feasible behaviors and the basis for exclusion of an individual behavior was not as important as the fact that it had been appropriately excluded.

Feasible Behaviors

Fifty-one behaviors were classified by the panel as feasible for assessment in an on-road test. Generally, these behaviors were rated highly overall by the reviewers or recommended for retention on the basis of panel discussion. The majority of the behaviors selected as feasible were from "Interacting with the Highway/Traffic Environment"; 33 feasible behaviors were from this category.

Noncritical Behaviors

Behaviors eliminated as noncritical were those determined by the panel to be either of no criticality to safety or of very low criticality. Examples of behaviors classified as noncritical are "shifts gears smoothly with manual shift," and "releases parking brake before moving."

Nontestable Behaviors

Behaviors classified as nontestable were of three types: Nontestable for Lack of Validity, Nontestable for Lack of Reliability, and Nontestable by Observation.

Nontestable for Lack of Validity

Behaviors were so classified if the panel believed that what would be observed in a test situation could not be assumed to correspond to real-world behavior. Examples are "wears glasses if necessary," "has not taken drugs or alcohol," and "adjusts mirrors."

Nontestable for Lack of Reliability

Behaviors were so classified if the panel believed that tasks calling for the behavior would not arise frequently enough to permit reliable measurement. Examples are "searches from side to side at a railroad crossing," and "slows for gravel and debris."

Nontestable by Observation

The panel classified behaviors as nontestable by observation if they could not be observed in an on-road test situation. Many search behaviors were placed in this group. Examples are "searches a suitable distance ahead," "searches the roadway ahead for warning signs," and "recognizes road users who are distracted."

Nonacceptable Behaviors

Behaviors classified as nonacceptable were of two types: Nonacceptable for Lack of Legal Requirement and Nonacceptable for Reasons of Safety.

Nonacceptable for Lack of Legal Requirement

This group of behaviors consisted of those that are not required by law. Examples are "fastens safety belts," "adjusts seat," and "locks door."

Nonacceptable for Reasons of Safety

Behaviors so classified by the panel were those requiring the creation of tasks that would pose risks to the driver, the examiner, the vehicle, and other road users. Examples are "responds to conflicts presented by intersecting vehicles," "responds to steering failure," and "recovers from skid."

Summary

Of the original list of 165 behaviors, none in the category of Driver/Vehicle Readiness was determined to be feasible for inclusion in an on-road test. Most were eliminated because test performance could not be assumed to reflect real-world behavior or because they were not required by law.

Four behaviors in the category of Vehicle Control were considered feasible. The rest were eliminated because they were thought to be noncritical to safety. Thirteen feasible behaviors were drawn from the category of Vehicle Maneuvering and 33 from the category of Interacting with the Highway/Traffic Environment. Only one feasible behavior was taken from Interacting with Highway/Traffic Hazards. A principal reason for exclusion of behaviors in that category was concern for safety.

A list of the 51 behaviors classified by the panel as feasible is provided below. (The classification assigned by the panel to each of the 165 behaviors is noted in the list provided in Appendix A.)

51 FEASIBLE BEHAVIORS

SKILLS

Vehicle Control

- Accelerates smoothly and evenly
- Brakes smoothly and evenly
- Releases accelerator while braking
- Grips steering wheel correctly

Vehicle Maneuvering

- Achieves and maintains appropriate speed moving forward
- Achieves and maintains appropriate speed entering curve
- Stops smoothly on level surface at designated point
- Maintains straight path moving forward
- Maintains curved path moving forward
- Judges clearance--fixed objects
    - Judges clearance between two objects
    - Judges clearance from object on left
    - Judges clearance from object on right
- Selects proper gap--moving objects
    - Entering roadway
    - Crossing traffic
    - Turning right into traffic
    - Turning left into traffic
    - Turning left across traffic

Interacting with Highway/Traffic Hazards

- Executes responses to emergencies
    - Brakes quickly

PRACTICES

Interacting with the Highway/Traffic Environment

- Searches ahead
    - Distance scanning
        - Looks from side to side
- Searches behind
    - Changing lanes
    - Merging
    - Turning
    - Stopping
    - Backing up

Searches rear quarter
- Entering roadway
- Changing lanes

Searches from side to side
- At blind intersections when privileged
- At normal intersections when privileged
- At intersections when burdened

Communicates intentions
- When turning
- When passing
- When changing lanes
- When entering or leaving traffic
- When merging

Manages speed--sight distance limitations
- Slows for curves
- Slows for blind intersections

Manages speed--signs, signals, markings
- Does not exceed posted speed limit, straight
- Does not exceed posted speed limit, curve/ramp

Manages space
- Avoids stopping in crosswalk
- Avoids driving in restricted areas
- Avoids crossing or straddling painted lines
- Selects appropriate travel lanes when turning
- Selects appropriate travel lanes when traveling through
- Maintains appropriate following distance when moving
- Does not slow unnecessarily

Manages space to the sides
- Maximizes separation from oncoming vehicle
- Maximizes separation from oncoming left-turning vehicle
- Avoids operating alongside vehicles
- Avoids operating in vehicles' blind spots
- Maximizes separation from parked vehicles

Obeys traffic signs and signals

STUDY OF MEASUREMENT METHODS

PURPOSE OF THE STUDY

Prior to the preparation of a preliminary version of the ADOPT, a study of measurement methods was conducted at the Texas Transportation Institute. The work at TTI had several objectives relating to measurement:

- o Establishment of Standards of Performance
- o Assessment of Methods of Measurement
- o Determination of Reliability of Measurement

This effort constituted Phase 2 of the project.

Standards of Performance

Before project staff could begin developing the ADOPT, it was necessary to establish acceptable standards of performance for many of the behaviors that comprise vehicle operating Skills.

Most operating Practices define a safe behavior that must be carried out as prescribed. For instance, obeying stop signs and using mirrors have clear-cut standards: The driver must stop and he must use his mirrors. For Skills, however, standards of performance are not as obvious. In the case of "accelerates smoothly and evenly," what determines that acceleration is smooth and even? At what point does acceleration become rough and uneven?

No one had ever defined what is "safe" with regard to operating Skills. They could be generally associated with safety, but no basis for determining unsafe levels of performance had been established. Thus, one objective of the study was to determine whether criteria of safe performance could be established for Skills and, if so, to identify those criteria.

Methods of Measurement

Reliable measurement of test behaviors requires accurate observation and recording of driver performance. During an on-road test, the examiner invariably sits in the front passenger seat right next to the applicant. In that position, the examiner is able to observe without difficulty much of the driver's performance. But some driving behaviors are extremely subtle--evidenced by slight eye movement or in the response of the vehicle to the behavior. Some can be sensed only through instrumentation. For this reason, an objective of the measurement study was to determine whether the responses of drivers can be observed and recorded with sufficient accuracy to permit reliable measurement and to establish the best methods of measurement.

Reliability of Measurement

Many of the candidate behaviors are responses to situations created by certain traffic and road conditions. If the situations do not occur often and uniformly enough, the corresponding behaviors cannot be reliably measured. Thus, an objective of the study was to determine whether situations calling for these behaviors could be encountered or created frequently enough to permit reliable measurement in a test. Beyond that, it was necessary to determine whether novice drivers make a sufficient number of errors to warrant measurement of the behaviors. If there were no variability in performance between novices and experienced drivers, measurement would reveal nothing. This aspect of the study did not require instrumentation and could have been carried out later during the pilot test. Nevertheless, the necessary data were readily available through the instrumentation, and use of these data greatly enhanced the pilot test version of the ADOPT.

MEASUREMENT PROCEDURES

Instrumentation

During the study, the Driver Performance Measurement and Analysis System (DPMAS) was employed. This system consists of a vehicle equipped with devices that permit a variety of measurements to be taken and recorded if desired. Measurement by human observation alone was not appropriate because:

- The rate at which situations and behaviors would occur would be too rapid to permit accurate recording.
- Many of the behaviors involved driver-control interactions that could be assessed only by means of instruments.
- The incidence of many behaviors could not be predicted sufficiently well in advance to permit the precoding necessary for manual recording of behavior. The observer would have had to watch for so many variables that manual recording would have been inaccurate or impossible.

The DPMAS is equipped with the following:

Vehicle Control Measurement Devices--Devices mounted on all vehicle controls that sense the magnitude of control input.

Vehicle Response Measurement Devices--Devices that sense vehicle response during motion. For the study, speed and lateral and longitudinal acceleration were selected for measurement.

Videotape Recorder--A recorder that records the images from three cameras. During the study, one camera was aimed at the road ahead, one was aimed at the road to the driver's left, and one was aimed at the driver to show driver eye and head movements. When the tape was played back, the view of the road ahead was displayed across the top of the screen, the view of the left side was at bottom left, and the view of the driver's movements was at bottom right. In addition, up to four items of information can be displayed in the upper left corner of the screen. For this study, the following were chosen: vehicle speed, longitudinal acceleration, lateral acceleration, and specific events codes. The specific events codes were two-digit numbers preassigned to each maneuver. The observer entered the number into the DPMAS by means of a keyboard.

The DPMAS is also equipped with a magnetic tape recorder that can record the vehicle control and response measurements. This equipment was not used for the study primarily because of the great difficulty of correlating the measurements with the tasks to which responses were made. Therefore, the most important measurements were selected for video display, rather than being recorded on magnetic tape.

Manual Observation and Recording

During the "test" portion of the on-street testing, signaling was recorded manually. Manual observation and recording was also used during the parking maneuvers to measure and record distance, time for completion, direction changes, and cone strikes.

STUDY METHODS

Both off-street and on-street testing were conducted.

Off-Street Testing

Off-street testing permitted assessment of operating Skills in a uniform set of maneuvers, free of traffic interruptions. Off-street testing was in two parts: Maneuvers on a closed course and handling exercises.

Closed Course

The closed course was one mile in length. Each subject traveled the course four times, twice in each direction. On each circuit, subjects were required to stop and start in a straight line, make turns at intersections, and negotiate curves.

Handling Exercises

Following the fourth circuit of the closed course, subjects performed three vehicle handling exercises: parallel parking, angle parking, and driving a lane of diminishing clearance. The parking maneuvers were created specifically to test specific basic vehicle handling skills, and not as a test of parking skill, per se. In the panel deliberations, parallel parking was eliminated from the list of candidate behaviors because it is not a critical driver behavior--it is not associated with high accident losses. Yet parallel parking and the other parking maneuvers are comprised by subsets of behaviors that were part of the candidate list. Therefore, the parking maneuvers were used for the study because they provided an efficient means of testing individual vehicle handling skills. The diminishing clearance lane, an alley created by pairs of traffic cones that are set progressively closer together, was designed to test judgment of clearance.

On-Street Testing

On-street testing permitted assessment of Skills and Practices in response to normal highway/traffic situations. The on-street testing involved both "test" and "free" operation.

"Test" Operation

The test routes was 19 miles in length. "Test" operation occupied the first 14 miles of the route. Subjects were asked to operate as safely as possible and were told that their performance would be recorded. The test route required subjects to perform such maneuvers as entering and leaving traffic, turning, responding to traffic signals, and negotiating curves.

"Free" Operation

After driving the on-street route for 14 miles, subjects were told that the test had ended and that they were to proceed back to the starting point. While manual recording of the data was terminated, the instrumentation and videotape recorder were left on. The purpose of this "free" operation was to record behavior under conditions in which subjects thought they were not being observed. Comparison of "test" and "free" operation could then be made to provide some indication of the validity of test performances as predictors of real-world driving behavior.

Test Population

The test population consisted of ten experienced drivers and ten driver education students. The latter served as surrogate "license applicants" and were tested to determine the likely error rate for candidate behaviors. The experienced drivers were tested in order to establish standards of acceptable performance. Comparison of the two groups was made to determine which behaviors would reveal poor performance and to establish criteria of deficiencies.

Behaviors Tested

The measurement procedures described permitted 14 of the 51 candidate behaviors to be assessed. Behaviors tested are listed below. The letters following the behaviors indicate where they were assessed. "R" indicates testing was off-street on the range. "S" indicates testing was on-street. "P-P," "P-A," and "P-D" indicate the behaviors were tested during parallel parking, angle parking, and parking in a lane of diminishing clearance, respectively.

Behaviors Tested

Skills

Vehicle Control

Accelerates smoothly and evenly R,S
Brakes smoothly and evenly R,S
Releases accelerator while braking R,S

Vehicle Maneuvering

Achieves and maintains appropriate speed entering curve R,S
Stops smoothly on level surface at designated point R
Judges clearance--fixed objects
- Judges clearance between two fixed objects P-D
- Judges clearance from fixed object on left P-A
- Judges clearance from fixed object on right P-A, P-P

Practices

Interacting with the Highway/Traffic Environment

Searches behind when changing lanes, merging, turning, etc. S
Searches rear quarter when entering roadway, changing lanes S
Searches from side to side at intersections S
Communicates intentions when turning, passing, etc. S
Slows for curves R,S
Obeys traffic signs and signals S

Behaviors Not Assessed

Available instrumentation did not permit certain candidate behaviors to be assessed during the measurement study. Some of these were Practices relating to proper management of space and speed and Skills involved in proper gap selection. Study of the behaviors not assessed had to be deferred until the pilot test of the preliminary ADOPT, when they would be assessed during actual administration of the test.

RESULTS

Skills

Acceleration

The novice driver's lack of precise foot control frequently evidences itself in sharp, "jerky" acceleration. This deficiency creates discomfort for passengers and some small danger to other road users. Beyond that, it is a symptom of a more general vehicle handling deficiency that is potentially hazardous. To arrive at an objective, defensible definition of acceptable limits, the acceleration of subjects from various stopping points on the closed off-street course and along the on-street test route was measured and the performances of novices and experienced drivers were compared.

While the experienced drivers had a somewhat higher mean rate of acceleration, there was no acceleration situation that reliably separated the two groups. Of the 40 accelerations each group performed on the off-street course, one novice acceleration reached .23 g, while none of the experienced drivers' accelerations exceeded .2 g. Accelerations in the .2 g range cannot be considered abrupt, however, and both groups exhibited accelerations as high as .29 g during the on-street testing.

It appears that the small variablility in acceleration observed during the study reflects differences in style rather than differences in vehicle handling skill. It is possible that a manual shift (the DPMAS has an automatic transmission) might have yielded different results. Nevertheless, since the overwhelming majority of cars used by license applicants have automatic shifts, a measure that applied only to manual shifts would have little utility for an on-road test.

Braking

In the off-street testing, no abrupt stops symptomatic of braking skill deficiency were observed. No negative accelerations in excess of .3 g were recorded. The six highest measured negative accelerations--ranging from .24 g. to .27 g--were performed by experienced drivers.

In the on-street testing, however, an entirely different picture appeared. Five accelerations in excess of .35 g were observed, three in "test" operation and two in "free" operation. All were made by novice drivers. Apparently when their attention was divided between vehicle handling and responding to highway/traffic conditions, braking skill deficiencies not evidenced in the off-street test were revealed. A negative acceleration of .3 g discriminated between novices and experienced drivers.

Speed on Curves

New drivers frequently lack the perceptual skill to judge the maximum rate at which a cruve can be safely negotiated. They may enter too fast or accelerate too rapidly. The result is a lateral acceleration that is uncomfortable and, under certain surface conditions, can be hazardous.

Six lateral accelerations of .4 g or more were measured, five on the off-street course and one on the on-street course. All were made by novices. Four of these accelerations occurred in 90° turns at the end of a straightaway. The other two involved 180° turns of large radii, also at the end of a straightaway. It did not appear that either the radius or the angle of the curve alone produced the excess acceleration, but rather these variables in combination with the approach configuration. In any case, a lateral acceleration of .4 g discriminated between novices and experienced drivers.

Braking in Curves

The potential hazard of lateral accelerations in curves is increased when accompanied by braking. Should the rear wheels lock at a point of high lateral acceleration, the vehicle could "spin out." Novices were expected to misjudge curves and have to apply the brake more often than experienced drivers. To find out if this were true, the position of the brake was analyzed at the point of maximum lateral acceleration on all curves on the off-street course. Brake application at the point of maximum lateral acceleration was most likely to occur under the same conditions as high lateral accelerations, that is, in 90° or 180° turns at the end of a straightaway. Right and left turns gave roughly equivalent results. Novices applied the brake during 31 percent of maximum lateral accelerations, while experienced drivers applied the brake in only 8 percent. Thus, application of the brake at the point of maximum lateral acceleration appeared to discriminate between novices and experienced drivers.

The potential hazard of brake application in a curve is proportional to the force with which the brake is applied. While brake force is measured by the DPMAS, it was not among the inputs recorded during the study. Nevertheless, one would expect brake force to reveal itself in the magnitude of the vehicle's negative acceleration at any point. To see whether brake force discriminated between novices and experienced drivers, the longitudinal acceleration at the point of maximum lateral acceleration was analyzed. There were no differences between the two groups.

The results of the two analyses indicted that while novices are far more likely to apply their brakes in a curve, it may be more of a precaution than an attempt to slow the vehicle. It probably reflects hesitancy or lack of confidence, rather than lack of skill. This uncertainty was also shown in duration of brake application, which was 60 percent longer among the novices than among the experienced drivers.

Stopping at a Designated Point

Stopping at designated points (designated by white lines on the pavement) was part of the off-street testing. Experienced drivers stopped much closer to the line than the inexperienced drivers. The mean for the experienced drivers was 2 feet short of the line, while for the novices it was 5 feet short of the line. No experienced driver stopped more than 7 feet from the stop line, while 30 percent of the novices did so. The results suggest that novices had more difficulty judging the position of the front of the vehicle than the experienced drivers. Novices apparently compensated for this difficulty by stopping farther from the line.

Duration of brake application leading up to each stop was also recorded. The novices applied the brake an average of 25 percent longer than the experienced drivers. Like the results for braking in curves, this suggests a deficiency in judgment accompanied by added caution.

While distance from the stop line and duration of brake application appeared to discriminate between novices and experienced drivers, the deficiencies of the novices hardly constitute hazards. No danger results from stopping well behind a stop line or from braking early. The judgment of stop line distance is a symptom of a deficiency in perceptual skill that could be hazardous under certain circumstances. Early braking, however, seems to indicate lack of confidence more than anything else.

Clearance Judgment--Fixed Objects

Parallel Parking

Subjects were assessed on the basis of the number of direction changes made during the maneuver, the number of cone strikes, the total time to complete the maneuver, and their final distance from the curb. Novice performances were compared with the performances of the experienced drivers. Analysis of results indicated the criteria of deficiency listed in Table 2 below. The number of subjects exceeding each criterion also is shown.

TABLE 2
CRITERIA OF SKILL DEFICIENCY FOR PARALLEL PARKING MEASURES

Measure	Criterion of Deficiency	Number of Measurements Exceeding Criterion: Novice	Experienced
Direction Changes	More than four	2	0
Cone Strikes	One or more	5	1
Total Time	More than 1 minute	3	1
Distance from Curb	More than 1 foot	6	3
	Total Measurements	16	5

On the basis of the criteria shown above, the parallel parking exercise separates novices from experienced drivers. In the case of Distance from Curb, a 2-foot distance distinguished more clearly between novice and experienced drivers (3 vs. 0). However, since the legal limit in most States is 1 foot, this distance was adopted as the criterion.

Angle Parking

The angle parking maneuver generally failed to show significant differences between novices and experienced drivers. Performance was measured in the same way as performance in parallel parking, except that cone strikes during exit from the parking space were also measured. Only in this last measurement was there a difference between the two groups. Cones were struck by eight of the novices but only three of the experienced drivers.

Parking in a Lane of Diminishing Clearance

There were no real differences between novices and experienced drivers in this parking maneuver. Six subjects in each group correctly perceived how far they could proceed down the lane. Cones were struck by two of the novices and one of the experienced drivers--certainly not a significant difference.

Role of Skills in Safety

None of the criteria for Skills involves either safety or legality directly. Their use in a license test would be based upon the general requirement that drivers be able to handle their vehicles. Drivers who cannot are a danger for three reasons:

1. With sufficiently long exposure, they may encounter situations in which their relative lack of handling skills will result in an accident.

2. Because their vehicle handling skills are not thoroughly mastered, they must direct some of their attention away from safe operating practices.

3. In an emergency situations, the ability to handle a vehicle often determines whether or not an accident will be avoided. Novices who can survive with poor handling skills during everyday driving are less well equipped to cope with emergencies.

A test that can identify those deficient in Skills and delay their licensing until they are qualified would appear to have accident prevention value. Drivers operating on a learner's permit are likely to confine their driving to areas in which they can operate safely despite their skill limitations. In addition, they must be accompanied by a licensed driver who can help compensate for their inability to devote as much attention to safe operating practices as do more proficient drivers.

Practices

Signaling

Use of turn signals was recorded manually and, therefore, only during the "test" portion of the on-street testing. (To have recorded signaling during "free" operation would have indicated to the driver that measurement was actually continuing.)

There were ten points on the street route at which signaling was required. The experienced drivers signaled properly 82 percent of the time, while the novices did so only 64 percent of the time. These results were consistent with earlier research indicating that when drivers know what they are supposed to do (signaling is required by law) and are motivated under test conditions to do it, those who are most proficient in vehicle handling skills will exhibit superior performance in safe operating practices. This has been attributed to the ability of those whose vehicle handling skills have become routine to devote more of their attention to following safe operating practices. According to this theory, safe operating practices are indirect measures of vehicle handling skills.

Visual Search

The subjects' observation of other vehicles through use of the mirror and by looking over the shoulder was assessed during lane changes and when entering traffic. Searching from side to side was assessed at intersections. The percentage of correct performance was tabulated. Results for the two groups are presented below.

TABLE 3
PERCENTAGE OF CORRECT PERFORMANCE
VISUAL SEARCH

Type of Operation	Novices	Experienced
Test	79%	67%
Free	78%	59%
Total	79%	65%

The experienced drivers were clearly surpassed by the novices. The major shortcoming of the experienced drivers was failure to look from side to side and over the shoulder properly; only 47% of these performances were correct. In the use of the mirror, experienced drivers did surpass the novices, 92% to 88%.

The fact that novices outperformed the experienced drivers is somewhat surprising in the light of the possible correlation between vehicle handling skills and safe operating practices, noted above in the discussion of signaling. The performance of experienced drivers may reflect the fact that they were less familiar with defensive driving practices than the novices, who were just completing driver education. In any case, the error rate among novices is sufficiently high to justify including visual search in an on-road performance test.

The failure of either group to show appreciable performance degradation in visual search during "free" operation suggests that the subjects did not relax their guard when they thought they were not being observed. This view is further supported by a correlation of .44 between performance under the two conditions for the combined groups. For analytic purposes, the two conditions can be viewed simply as proportions of the same test. A test-retest correlation of .44 is fairly high for a single measure.

Other Behaviors

The performances of both novices and experienced drivers in slowing for curves and obeying traffic signs and signals were virtually errorless, and no significant differences were noted between the two groups.

Summary

The results of the study at the Texas Transportation Institute suggested that the following behaviors and criteria of deficiency were appropriate for use in assessing Skills:

Braking

- Negative accleration of .3 g or greater

Speed on Curves

- Lateral acceleration of .4 g or greater
- Brake application at point of maximum lateral acceleration

Stopping at Designated Point

- Stopping more than 7 feet from designated point

Clearance Judgment--Fixed Objects (Parallel Parking)

- More than four direction changes
- One or more cone strikes
- More than 1 minute to complete maneuver
- Final distance from curb more than 1 foot

Turning to Practices, the analyses of Visual Search and Signaling behavior reveal error rates that are sufficiently high to warrant inclusion of these performances in an on-street test. Differences between novices and experienced drivers suggest that signaling may be an indirect measure of skill. This does not seem to be true of visual search, possibly because many experienced drivers are not aware of its importance.

DEVELOPMENT OF THE ADOPT

Following the Phase 2 study of measurement methods, Phase 3 was initiated with development of the ADOPT. This section describes the preliminary version of the ADOPT that was developed. The two remaining Phase 3 activities, the Pilot Test and Field Test of the ADOPT are described in the succeeding sections.

ADOPT Test Behaviors

Using the list of 51 behaviors selected by the Selection Panel, project staff prepared a preliminary ADOPT for pilot testing. Two of the behaviors--slowing for curves and obeying traffic signs and signals--were eliminated because of almost errorless performance by drivers in the Phase 2 study. The other behaviors were retained for further study in the pilot test of the ADOPT. The numerical criteria for vehicle handling skills derived from the analyses of Phase 2 data were applied to the scoring of individual behaviors in that category.

At this point in the development of the ADOPT, the nomenclature for some of the behaviors was changed. This was done primarily to permit all behaviors to be described in no more than three words, for the sake of economy and to simplify reference to the behaviors. Additionally, while all of the behaviors on the list were incorporated into the test, similar behaviors were grouped together under a common term. For example, gap selection behaviors involving moving objects were combined under the term "Gap Selection."

Behaviors included on the preliminary ADOPT are listed below, along with brief descriptions of the required performances.

Handling--Must maneuver the vehicle along a prescribed path, both forward and backward, in tight quarters. (Performance can be assessed off the street on a specially created course or on the street along a curb.)

Brake Application--Must regulate brake pressure in order to stop smoothly, at a force of deceleration less than .3 g.

Wheel Grip--Must operate with both hands on the wheel.

Controlled Stop--Must bring the vehicle to a complete stop at or near a designated point.

Rapid Stop--Must bring the vehicle to a stop as quickly as possible without locking the wheels.

Lanekeeping

Straight--Must operate within the bounds of proper travel lanes.

Curve--Must operate within the bounds of proper travel lanes while negotiating the entire length of a curve.

Turn--Must operate within the bounds of proper travel lanes while negotiating a right or left turn at an intersection.

Gap Selection--Must make the correct decision in response to the first gap for (1) traversing, turning left or turning right in the path of cross traffic or (2) making a left turn across oncoming traffic.

Speed Management

Selecting Speed: Intersection--Must select a speed that will allow observing for traffic on the cross street.

Selecting Speed: Curve--Must adjust speed to negotiate a curve or turning with a lateral acceleration force of less than .4 g.

Selecting Speed: Limit--Must adjust to any new speed limit and must not exceed the posted limit by more than 5 mph.

Maintaining Speed--Must not drive at more than 10 mph below the posted limit.

Observing

Behind--Must check the inside rear-view mirror to determine the presence and distance of other road users.

Rear Quarter--Must make both mirror and over-the-shoulder checks to determine the presence of traffic before initiating a right or left lane change or merge.

Side--Must observe to the sides in the direction of cross traffic at an intersection.

Communicating

Lane Change--Must signal before entering a new lane on the roadway.

Turn--Must meet legal standards for signaling when turning or leaving the roadway.

Distance Management

Adjacent Vehicle--Must maintain adequate separation from traffic to the side, as well as parked vehicles, pedestrians, and vehicles likely to enter the roadway.

Oncoming Vehicle--Must maintain adequate separation from oncoming vehicles.

Following Vehicle--Must maintain adequate separation from the vehicle ahead.

Restricted Travel--Must comply with travel restrictions as indicated by signs or roadway markings.

Administrative Procedures

Once the nomenclature for test behaviors had been established and the performances described, administrative procedures, including procedures for selection of test routes and situations, were also developed. Methods for implementing, administering, and scoring the ADOPT were documented in two manuals: one for administrators and one for examiners. These manuals provided all information that would be required by a State DMV to implement the ADOPT. Administrative procedures remained constant through subsequent versions of the ADOPT.

Administration Time

The preliminary ADOPT was designed to require approximately ten minutes for administration. This amount of time represents the maximum allocated by any State to administration of road tests to applicants for a regular driver's license. Because of the one-to-one examiner:applicant ratio, administration of the road test is by far the most expensive aspect of the licensing operation. An increase of even five minutes in testing time would have a significant impact upon the operating costs of licensing agencies. It is the belief of those licensing officials with whom administration time was discussed during test development tnat a road test requiring more than ten minutes to administer would not be implemented as a regular licensing examination.

Personnel

The test was designed to be administered by the type of people who currently administer road tests. This means individuals of average intelligence having no educational background in traffic safety, driver behavior, or any other academic discipline related to driver testing.

While administration of an examiner training program could be anticipated, it could not require more than two days. Turnover among examiners is characteristically fairly high and new examiners are continually being trained. A program requiring any more than two days to administer would end up diverting too much personnel time to the giving and receiving of instruction.

Equipment

The need for equipment to support administration of the test had to be kept to a minimum, in order to avoid adding significant cost items to administration of the test. Experience with the Motorcycle Operator Skill Test (MOST) clearly indicated that equipment requirements pose an obstacle to implementation of a license test. Any equipment that is required must be highly portable, readily installed, and easy to use.

During training, an accelerometer is used to assist examiners in obtaining reliable estimates of vehicle acceleration, including:

> Negative acceleration--Negative acceleration must be less than .3 g in Brake Application during a stop.
>
> Lateral acceleration--Lateral acceleration must be less than .4 g in Selecting Speed: Curve.

Originally, all examiners were to have been provided an accelerometer to be used during test administration. However, the appropriateness of the device was questioned by licensing officials because of both the cost of providing each examiner with an accelerometer and concern about the possibility that the equipment aid might undermine the applicant's confidence in the examiner's ability to assess qualifications. The need for the device disappeared when it became clear during the training that examiners could be trained to make very reliable judgments.

Test Situations

The procedure under which the ADOPT is administered calls for examiners to observe and assess specific performances in specified situations. This procedure is based upon prior research indicating that reliable assessment of applicant performance can be made only when the examiner's attention is directed at specific behaviors.

Different examiners will tend to notice different things. One may tend to notice errors while another will notice both errors and correct performances. Attention of some examiners will be toward one type of behavior, others toward a different type of behavior, depending upon what they happen to believe is most important. The result is often that the applicant's score on the license test depends as much upon what the examiner happens to notice as it does upon what the applicant actually does.

To assure the reliable assessment of driver performance, tests developed more recently have employed the practice described a moment ago, that of calling upon examiners to assess specific behaviors in response to specific situations. The Safe Peformance Test, Michigan Road Test, and Motorcyclist In-Traffic Test, all referenced earlier, employ this practice.

Test situations fall into two categories:

> Roadway situations, created by characteristics of the roadway such as intersections or curves.

Traffic situations, created by characteristics of traffic, e.g., a vehicle ahead (following distance) or cross traffic (gap selection).

For each performance check, the roadway and traffic situations needed to give rise to the performance to be checked have been identified. These situations may be found in both the Administrator's Manual and Examiner's Manual for the ADOPT.

Test Locations

To assist the examiner, locations along the route in which the situations are most likely to arise are selected in advance and the specific behaviors to be assessed are identified. Upon reaching a location, the examiner need only attend to the specific behaviors selected for assessment at that location.

Care must be taken in selecting locations to assure that a representative sample of applicant performance is attained. Applicants are not totally consistent; they can respond to a situation correctly at one location and incorrectly at another. If most of the applicant errors occur at locations where the behavior is not being scored, the result is an invalid assessment of performance, and a frustrating experience for the examiner.

Particular care must be taken in selecting locations for traffic situations. There is, of course, no way of assuring that the traffic conditions that make up a test situation will prevail at the time the applicant reaches the particular location. In developing the ADOPT, a great deal of attention was devoted to characteristics of locations at which the various required traffic situations are most likely to arise.

To provide a reasonably high likelihood that the traffic conditions making up a test situation will occur at a location, traffic situations are checked within "zones" ranging from a block or two in the city to as much as a quarter of a mile in the open road. If the stretch of road over which a test situation is likely to arise exceeds these limitations, two or more zones can be strung together. (Since there is room on the test form for scoring only one instance of behavior within a zone, it is better to have several short zones than one long zone.)

Selection of Performance Checks

The performances that are checked by the preliminary ADOPT were described earlier in the discussion of Test Behaviors. Each individual assessment of a behavior is referred to as a "Performance Check." After appropriate test situations have been located, the specific performances to be checked at each location are selected. A particular location may provide the situations required for several different performance checks. In order to provide the most representative possible example of behavior, the selection of performance checks for the ADOPT is guided by the following considerations:

Type of Situation--Situations that depend upon traffic conditions must be sampled more frequently than those depending upon roadway conditions, since the required traffic conditions may not prevail at the selected location eacn time the test is administered. For example, several checks of Gap Acceptance must be called for to be reasonably sure of providing the required traffic conditions even once. However, the conditions needed for testing use of turn signals will be present at any intersection where a turn is made.

Type of Performance--Performances that represent day-to-day practices appear to be less stable than those tnat reflect skill and therefore must be sampled more frequently. For example, a driver may signal one turn and neglect his signal on another. Or, he may check his mirror before making one change and not do so on another. On the other hand, a driver who has the skill to remain within a lane in one curve will proably do so on almost all curves. Because of the stability of skills, only one check of Handling and Rapid Stop is called for in the ADOPT.

Number of Checks--The ADOPT allows as many as four performance checks to be made at any one location. Since the applicant's performance cannot be observed and recorded at the same time, the examiner must wait until the location is passed before recording responses. Four represents the upper limit of performances that can be reliably observed and remembered. The observations required of the examiner must be compatible. For example, the examiner cannot be asked to check driver eye movement and vehicle position at the same time.

Separation of Points and Zones--On the test route, performances checked at "points" must be separated from those checked in "zones." Under the ADOPT procedure, for example, the examiner would not be called upon to check to see if an applicant checked a blind intersection if that intersection were in a zone where lane-keeping was being checked. It proved too difficult for examiners to remember to make a check of performance at a specific point in a zone in which their attention was directed to other performances.

Scoring System

The ADOPT scoring system involves scoring tne results of individual performance checks and then combining those results to obtain a total test score.

Scoring Individual Checks

The ADOPT employs a highly objective system for scoring performance. For each performance check, there is a set of criteria that define what is acceptable. The examiner is called upon simply to judge whether or not the applicant's performance has met those criteria. This system can be contrasted with a more subjective approach in which applicants are rated on how well they handle an entire task, e.g., "good," "excellent," or "5" on a scale of 1-10.

The objective approach employed in the ADOPT has two major advantages. The first is that it leads to high examiner reliability. The examiners are more likely to agree upon whether an applicant did or did not look in the left outside mirror before initiating a lane change than they are on how "good" the lane change was. High examiner reliability is a necessary condition for validity--how can a score be considered valid if two observers disagree as to what the score should be? It is also a necessary ingredient of a test that is equitable--how can a test be fair if different examiners score the same performance differently?

Obtaining a Total Score

An applicant's score on the ADOPT is the percentage of test situations responded to correctly. It is obtained by counting the number of correct responses and dividing it by the total number of responses correct and incorrect.

If applicants had an opportunity to respond to all of the situations on a test, it would only be necessary to add up the total number of correct responses in order to obtain a score. The number correct and the percentage correct would be proportional to one another and give the same results. However, not all of the test situations will arise during a given test administration, particularly those that are traffic dependent. Expressing results as a percentage controls for variations in test situations from one applicant to another.

If applicants were scored only in terms of correct responses, the applicant who encountered the greatest number of traffic situations would have a definite advantage. Conversely, if only errors were scored, the applicant who encountered the heaviest traffic would be at a definite disadvantage. The use of percentages adjusts for differences in traffic conditions and other test situations much in the same way as a batting average adjusts for differences in number of trips to the plate, or a fielding average adjusts for differences in number of fielding chances.

Unscorable Situations

If the situations needed for a performance check do not arise, the performance cannot be scored. Because of the system of obtaining a total score, a failure of a situation to arise does not affect an applicant's score favorably or adversely.

Where a test situation does not arise, it could simply be ignored. However, in the ADOPT, there is a scoring category labeled "Not Applicable," which is used to indicate when a performance cannot be scored. Use of the "Not Applicable" scoring category is intended to force examiners to respond overtly to each performance check, lessening the chances that the check might be overlooked.

Test Form

The ADOPT furnishes examiners a Test Form to guide them in administering the test. Specifically, the Test Form provides:

- o Directions to applicants
- o Locations of performance checks
- o Performances checked
- o A score sheet

A portion of a sample test form, taken from the ADOPT Field Test, appears in Appendix D.

Direction to Applicants

The examiner must provide directions to guide the applicant around the test route. In the ADOPT, these directions are provided on the Test Form so that they will be given in a uniform manner on each test administration. Where lengthier specific instructions are required to evoke a particular performance (e.g., Rapid Stop) these instructions are referenced on the Test Form. The use of written directions and instructions also fosters uniformity by discouraging casual remarks by the administrator.

A number of conventions are employed in giving instructions and directions in order to assure clarity of communication. These conventions are presented in the Examiner's Manual.

Locations

The Test Form identifies the locations at which performances are checked as well as those at which instructions and directions are to be given. In some road tests, locations are indicated on a map of the route. While a map is a useful device for acquainting examiners with a route, it does not provide a very suitable format for identifying performance checks or recording applicant performance. Moreover, its virtues in describing the route disappear after a few test administrations. A license examiner who administers a test over the same route many times a day certainly does not need a map.

In the ADOPT, locations are described as briefly as possible in order to avoid cluttering the Test Form. Intersections are located by identifying the street being driven and followed by the street being crossed or entered, e.g., "Main at Elm." Other points along the route are similarly identified, e.g., "12th Street at Bide-a-Wee Motel." Zones identified by the street being given followed by the beginning and end points, e.g., "Johnston between 6th and 8th."

Performances Checked

The Test Form identifies the specific performances to be checked at each location. A two-digit letter code is used as a space-saving method. Because the letter codes employ the first initial of the performance checks, they are very easy to learn and remember.

Score Sheet*

Next to each performance check is provided a space for recording the applicant's score. For each check, the examiner checks one of the following categories:

YES--The applicant's behavior met the specified criteria of acceptable performance.

NO--The applicant's behavior did not meet the specified criteria of performance.

Not Applicable--The performance could not be scored.

Upon completion of the test, the number of "YES" and "NO" checks is totaled. As described earlier, the total score is derived as follows:

$$\text{Total Score} = \frac{\text{TOTAL OF YES RESPONSES}}{\text{TOTAL OF YES + NO RESPONSES}}$$

* License examiners typically use "score sheet" to refer to the entire test form. For this reason, the test form is referred to as the "score sheet" in the Examiner's Manual and Administrator's Manual.

PILOT TEST

The preliminary version of the ADOPT containing the behaviors described in the preceding section was pilot tested in Oklahoma, with the cooperation of the Oklahoma Department of Public Safety (DPS). This section describes the purpose, methods, and results of the pilot test.

PURPOSE

The purpose of the pilot test was to assess the reliability and validity of the ADOPT as a whole and of its individual measures. The results obtained from this analysis were to be applied to revision of the ADOPT in order to improve overall reliability and validity.

Reliability

The overall reliability of a test is the extent to which individual administrations provide reliable samples of the total population of behaviors that the test is intended to measure. Three types of reliability were investigated in the pilot test of the ADOPT:

- Examiner Reliability--The extent to which the examiners accurately observe, collect, and report applicant behavior.
- Sampling Reliability--The extent to which individual administrations regardless of location provide representative samples of the total performance.
- Measurement Reliability--The extent to which test results are reproducible across examiners and locations. Measurement reliability is a function of both examiner reliability and sampling reliability.

Target standards of reliability for examiner and sampling reliability were set at .8 and .6 respectively. A .8 level of examiner reliability was obtained in the USC and MSU road tests. While these tests were somewhat longer than the ADOPT, the target seemed attainable. The sampling reliability target of .6 reflects the varying conditions found along test routes and the discrepancies in scores anticipated from different segments of a route. Correlations approaching .6 were found in pilot testing of the MIT, making the objective seemingly an attainable one. With these examiner and sampling reliability coefficients, an overall measurement reliability in the neighborhood of (.8 x .6 =) .5 was anticipated.

Validity

The pilot test also assessed the validity of the ADOPT--the degree to which the test is able to assess real-world driving behavior. The manner in which behaviors were selected for inclusion in the ADOPT endows it with a degree of content validity. Nevertheless, the behaviors that comprise the ADOPT consist of responses to a set of prescribed situations. The content validity, therefore, is limited to the sample of situations. Whether the behaviors that occur in response to those situations are representative of real-world behavior is something that cannot be assessed on the basis of content. If the ADOPT is truly valid, the responses as well as the situations must be representative of real-world behavior.

There is reason to be skeptical concerning the ability of test responses to provide samples of real-world driving behavior. In the case of vehicle handling responses, which are dependent primarily upon highly internalized, routinized skills, what participants do during the test probably reflects the skill with which they usually drive. In contrast, most of the responses related to safety of operation (e.g., signaling, following at a safe distance) are greatly dependent upon the driver's motivation to perform them. The motivation that prevails during a test is not likely to be representative of that characterizing the real world, and substantial differences between test performance and real-world behavior could be expected.

To support validity assessment, data on nontest driving behavior were collected on a subsample of applicants.

METHODS

Selection of Test Sites

Oklahoma was selected for the pilot test for two reasons. One was that the State was willing to provide personnel time at no cost to the contract. This included examiner time in administering the ADOPT, DPS staff time to assist project staff in route selection, and administrative and management time (e.g., for processing data). The second reason was that the Oklahoma DPS had gained experience in an earlier pilot test of an on-street performance test. This meant that required training time would be reduced and ensured that the DPS was fully aware of the general requirements of a pilot test.

It was decided to conduct the test at two sites within the State of Oklahoma: Oklahoma City and Tulsa. These sites were selected on the basis of the weekly volume of applicants. Applicants averaged 100 per week at the Tulsa licensing station, more than any other station in the State. At Oklahoma City, the average number of weekly applicants was about 50. These numbers were sufficient to meet sampling requirements for the pilot test within the allotted time.

Route Selection

Project staff selected two test routes in Tulsa. This was done because the high volume of applicants permitted the testing of two applicants at a time. Two routes also provided a wider sample of route conditions. Only one route was selected in Oklahoma City, since the volume of applicants there did not support the use of two routes.

Because the ADOPT is intended to serve ultimately as a valid, reliable test regardless of where it is administered, an attempt was made to select routes for the pilot test that incorporated urban, suburban, and rural environments. The two Tulsa routes each contained urban, suburban, and rural environments. The Oklahoma City route contained only urban and suburban segments because a rural segment of sufficient length was not available.

Each of the three routes was designed to provide a maximum variation in conditions. The route segments (urban, suburban, and rural) were contiguous, and provided enough exposure to traffic and roadway conditions to provide a broad sampling of constituent behaviors within the given environment.

The three test routes each required about 20 minutes of driving time. While the ADOPT itself is intended to require only 10 minutes of driving time, pilot test routes were lengthened to permit assessment of sampling reliability. Such an assessment required more than one sample of behavior for each applicant. For this reason, each applicant essentially drove two 10-minute routes.

Pilot Test Sample

The pilot test sample consisted of 150 subjects, evenly divided among the three routes. The subjects were all novice drivers applying for a license for the first time. Previous NPSRI research in the development of an on-road performance test indicates that measures of examiner reliability and sampling reliability tend to be unstable with samples of less than 100. Therefore, a sample size of 150 was chosen to help ensure accurate results.

Test Procedures

At each site, a coordinator assigned the first available applicant to the examiners who were responsible for testing. Each test was conducted by two examiners. One rode in the passenger seat in the front of the car, while the other rode in the rear center seat. They reversed positions prior to each successive road test. A plausible explanation for the presence of the second examiner was given to the applicant (e.g., on-the-job training).

Prior to the road test, each applicant was given a brief handout that explained the test contents and standards for performance. Examiners also provided a standard oral explanation of the road test.

As the applicant drove, each examiner used the ADOPT test form to score the applicant's performance. This was only for the purposes of the pilot test. Pass/fail determination was made by the examiners using regular Oklahoma DPS procedures.

Upon completion of the road test, applicants were informed whether they had passed or failed. As applicants left the licensing station, they were followed surreptitiously by a field assistant and their nontest driving behavior was recorded. Videotaping was done only of applicants at the Tulsa licensing station. The Tulsa site was selected for videotaping because of the larger number of subjects and the higher anticipated applicant flow rate. Videotaped data were collected on all applicants with the exception of those who had an adult passenger (a parent or other authority figure) whose presence would be likely to inhibit driving practices that are unsafe or illegal.

The target was 5 minutes of videotape per applicant. This was determined to be adequate for project purposes and also meant that the field assistant could return to the station in time to follow the next applicant. It also confined the observations to areas near the licensing station, providing similar situations across applicants, and kept the magnitude of the data reduction task within manageable limits. Actual taping averaged 7 minutes per applicant, largely because routes leading away from the Tulsa licensing station were such that it took the field assistant about 7 minutes to reach a point where he could turn around.

Videotaping was accomplished by means of a camera mounted in a fixed position on the passenger side in the front of the car. Fixed mounting eliminated the need for a separate camera operator and provided a steady image. The camera was pointed straight anead. Target vehicles left the field of view only during right angle turns at intersections. The camera zoom lens was adjusted so that the target vehicle filled approximately half the screen at normal following distance.

In addition to videotaping the applicants, the field assistant used the microphone on the camera to note certain behavior that perhaps would not be discerned from the videotape. Tape-recorded information included such things as:

Speed--If greater than 5 mph over or 10 mph under the speed limit.

Position--Position relative to crosswalks and stop lines; position during right-angle turns when applicant's vehicle was out of sight of the camera.

Visual Scanning--All head movements indicating visual scanning.

After the road tests, the two coordinators collected the test forms and reviewed them for completeness. At the end of each day, completed test forms were mailed to project staff.

Testing along a given route ceased as soon as the target of 50 applicants was achieved. The pilot test was completed in one week.

Pre-Pilot Test Preparation

Project staff made two site visits to Oklahoma to prepare for the pilot test. During the first visit, DPS personnel were briefed on the nature and purpose of the pilot test and on the ADOPT itself. Preliminary administrative procedures were drafted, routes were selected, and arrangements were made to secure personnel for the pilot test.

During the second site visit, all administrative procedures were made final, and project staff provided personnel training.

Personnel Training

Pilot test personnel consisted of six examiners (two who served as coordinators), and one field assistant. Project staff provided overall supervision and training.

Examiners

The examiners were responsible for conducting the road tests and scoring each applicant using the test forms provided by project staff. Two examiners were assigned as a team. Both rode with each applicant, each filling out a test form.

Project staff provided two days of examiner training. The first day involved instruction in the objectives of the ADOPT and the scoring procedures. The second day was devoted to practice testing. The pairs of examiners assigned to the three routes administered the ADOPT to surrogate applicants (members of the project staff) along their assigned routes. Practice continued until the two examiners on each route achieved an inter-examiner reliability level of .8.

Coordinators

The coordinators were DPS examiners and were responsible for issuing test forms to the examiners, assigning examiners to applicants, and mailing completed test forms to the project staff. Project staff met with the coordinators at the time of practice testing to review and explain their responsibilities.

Field Assistant

During the pilot test, two sources of data were accumulated: the examiner test forms and videotapes of the nontest behavior of applicants. Videotaping was the responsibility of the field assistant. This individual was engaged locally and was not connected with the DPS. The field assistant was briefed by project staff and performed several practice tapings to become familiar with pilot test procedures.

Analysis of Videotapes

Approximately four hours of videotape were accumulated on 35 applicants. The videotapes were both scored objectively and rated by a panel of experts.

Scoring

First, members of the project staff reviewed the videotapes and scored the drivers using objective criteria. Points were accumulated for incorrect or unsafe performance, so that a high score indicated poor performance. One point was accumulated for an error in any of the following performances:

Brake Application
Controlled Stop
Lanekeeping: Straight
Lanekeeping: Curve
Lanekeeping: Turn
Gap selection
Selecting Speed: Limit
Selecting Speed: Intersection
Selecting Speed: Curve
Maintaining Speed
Observing: Rear Quarter
Observing: Side
Communicating: Lane Change
Communicating: Turn
Distance: Adjacent Vehicle
Distance: Oncoming Vehicle
Distance: Following Vehicle
Restricted Travel

The objective scores for each subject videotaped were compiled by summing the errors in each individual category observed on the videotape. The scores in each category were accumulated to form separate totals for safe driving practices and vehicle handling skills.

Rating

Second, a six-member panel reviewed the videotapes. Applicant performance in the two behavior categories (vehicle handling skills and safe driving practices) was rated by the panel using the following scale:

1 = Applicant is operating very safely; areas of improvement are very difficult to identify.

2 = Applicant is operating safely, but there is some room for improvement.

3 = Applicant is operating fairly safely, but there is definitely room for improvement.

4 = Applicant is operating with marginal safety and needs a lot of improvement.

5 = Applicant is operating unsafely and should not be on the street.

Intercorrelation of panel ratings prior to discussion was obtained through interclass correlation analysis. The mean intraclass correlation coefficient, reflecting the amount of agreement between the panel of six raters, was .89 for vehicle handling skills and .91 for safe driving practices. These correlations indicate very high agreement among the panelists. While agreement does not assure accurate ratings of skill and practice, it is certainly encouraging.

Once the reliability of panelist ratings had been assessed, the videotape records of subjects having a spread of ratings of more than 2 points on the 5-point scale employed were reviewed and discussed until a consensus was reached. The ratings of the various panelists were then averaged to obtain overall ratings of skill and practice.

RESULTS

Data from the pilot test were analyzed to assess examiner reliability, sampling reliability, measurement reliability, and external validity of the interim ADOPT.

Classification of Behavior

The behaviors making up the ADOPT may be divided into two categories:

Skills--Those behaviors that are concerned primarily with controlling motion of the vehicle and dependent primarily upon perceptual and motor skills for successful performance.

Practices--Those behaviors that are concerned largely with assuring the safety of operation and which are dependent more upon motivation than skill.

The effect of the testing process upon measurement is likely to differ somewhat across these two behavior categories. The operating skills observed on a test are likely to be fairly similar to those that are seen in day-to-day operation. When it comes to safety of operation, however, what is revealed in a test may differ substantially from everyday practice. Because of these anticipated differences, it seemed desirable to separate the two categories of behavior in the analysis of data. The behaviors were classified as follows:

Skills	Practices
Brake Application	Wheel Grip
Controlled Stop	Selecting Speed:
Rapid Stop	Limit
Lanekeeping:	Intersection
Straight	Maintaining Speed
Curve	Observing:
Turn	Behind
Gap Selection	Quarter
Selecting Speed: Curve	Side
Handling:	Communicating:
Time	Lane Change
Direction	Turn
Strikes	Distance:
Observing	Adjacent
Position	Oncoming
	Following
	Restricted Travel

Examiner Reliability

Examiner reliability is the extent to which one examiner's score is representative of all examiners' scores. It was measured in two ways.

Constant error was examined by comparing the mean scores given by the two examiners assigned to each route. The results appear in Table 4 below.

TABLE 4
MEAN SCORES BY EXAMINER POSITION

	Front	Rear
Total	.77	.78
Skills	.76	.79
Practices	.78	.78
Urban	.81	.81
Rural	.69	.70
Suburban	.78	.80

In no case did the means differ by more than two percentage points. It would appear the scoring system eliminated any bias caused by the position of the examiner in the vehicle.

In addition to comparing examiners for overall test scores, comparisons between examiners were made for Skill and Practice subtests as well as each of the 25 behavior categories. Because of the relatively small number of observations involved, differences of as much as 5% were observed. Nevertheless, only one of the many differences was significant at the .05 level, and that is in itself a chance finding given the large number of comparisons made.

Variable error was measured by correlating the scores of front and rear examiners. The overall estimate of reliability was obtained by averaging the individual correlations obtained on each route, rather than by pooling across routes. Owing to differences in the level of performance among routes, pooling would have introduced a spurious correlation. The average total examiner reliability was .76, which is acceptable although slightly below the goal of .80.

Analysis of reliability by route location produced coefficients for urban, suburban and rural locations of .72, .74 and .64 respectively. These differences parallel differences in the number of performance checks made in each area. Ninety-two checks were made in urban areas, and 142 were conducted in suburban locations. Only 56 checks were made in rural areas since the third route (Oklahoma City) did not include a rural segment.

Analyzing examiner reliability by subtest, we find the average coefficient of correlation was .79 for practices, while only .61 for skills. Such a difference was anticipated. While scoring of practices generally requires a simple determination of whether or not a behavior occurred, assessment of skills typically requires a judgment of quantity, e.g., magnitude of acceleration, distance.

Sampling Reliability

Sampling reliability was determined by comparing the measurements of behavior from the three locations--urban, suburban, and rural. (On the two Tulsa routes, applicants were tested in each of the three locations. The Oklahoma City route did not contain any rural locations.)

Sampling reliability, like examiner reliability, was assessed in terms of constant and variable error. Constant error was assessed by comparing mean scores across all three locations. There was no sizable difference between the average scores for urban and suburban locations, which were .81 and .78 respectively. Scores in rural areas averaged .69, considerably lower than the other two. This difference, however, is totally due to one of the two routes. On the other, the rural mean was within two percentage points of the urban and suburban locations. We cannot say that rural locations are inherently given to "harder" performance checks--only that such was apparently the case for one route.

Variable error was estimated by intercorrelating the scores of applicants across three pairs of route locations. To control for any error introduced by the position of the examiner, the reliability estimates were produced separately for front and rear positions. The correlations were averaged across the three routes rather than pooled in order to avoid the spurious effects of differences in the level of performance between the routes. The results appear in Table 5 below

TABLE 5
SAMPLING RELIABILITY

	Front Examiner	Rear Examiner
Urban vs. Suburban	.54	.58
Rural vs. Suburban	.45	.44
Rural vs. Urban	.59	.49

None of the correlations reached the standard of .6 established beforehand. Nevertheless, the correlations between urban and suburban locations--the only ones affording a reasonably large sample of behavior--were fairly close to the standard. Further refinement of the ADOPT was expected to lead to achievement of the original goal.

Measurement Reliability

The total reliability of the measurement process was formulated by comparing the results obtained by one examiner on one route segment (or location) with the scores of the other examiner on another route segment.

All possible pairs of comparisons were averaged to arrive at the overall measurement reliability. Table 6 shows the comparisons that were made.

TABLE 6
MEASUREMENT RELIABILITY

Locations Compared	Correlation
Front Urban/Rear Suburban	.50
Front Suburban/Rear Urban	.40
Front Urban/Rear Rural	.39
Front Rural/Rear Urban	.48
Front Suburban/Rear Rural	.36
Front Rural/Rear Suburban	.41
Overall	.42

The overall average correlation was .42, which was lower than the goal of .5 but close enough to place the goal within striking distance.

Validity

The validity of the ADOPT was assessed by correlating nontest behavior with performance on the ADOPT. The correlations between ADOPT test scores and both scored and rated nontest behavior appear in the Table 7 below.

TABLE 7
CORRELATION OF TEST AND NONTEST BEHAVIOR

Nontest	Test	
	Skill	Practice
Scored		
Skill	-.41	.17
Practice	-.30	.10
Rated		
Skill	-.18	-.01
Practice	-.23	-.04

Skill Measures

Because of the way in which nontest behavior was scored and rated, negative correlations with the ADOPT test scores would be expected. The obtained correlations show that the ADOPT skill measures correlated with both skill and practice in real-world, nontest driving behavior. The correlation between test and nontest skill measures (-.41 and -.18) lends support to the hypothesis that the way people handle a vehicle on the ADOPT is somewhat similar to that which characterizes their real-world driving and that the ADOPT scale measures are predictors of skill-mediated driving behavior. The fact that the correlations are relatively low is not surprising in view of the limited opportunity to measure handling skills through the videotape performance.

The ADOPT skill measures also correlate with both scored and rated nontest practices (-.30 and -.23). The correlation between skill and practice on the ADOPT itself was .40. A correlation between skill and practice has been observed in testing of motorcycle riders.

The relationship between skill and practice has been explained by the need for a certain level of vehicle handling skill in order to be able to devote attention to use of safe operating practices. It has been hypothesized that most vehicle operators know what they are supposed to do. The less-skilled operators, because they must concentrate their attention on vehicle handling, tend to forget to put this knowledge into practice. The skillful drivers were able to concentrate more of their attention upon employing safe driving practices in the real world just as they did on the test.

Practice Measures

ADOPT measures of safe driving practice do not appear to be related to any aspect of real-world driving behavior. Their correlations with scored nontest behavior were in the wrong direction, while correlations with rated behavior were close to zero. This result supports the hypothesis suggested in the earlier discussion of validity that the use of safe driving practices in a test situation is not representative of normal driving behavior.

Overall Assessment of ADOPT

The results of the pilot test reflected favorably upon the ADOPT, given its state of development. Coefficients of interexaminer, sampling, and measurement reliability were very close to the standards set in advance. Improvements in the test between pilot test and field test were expected to allow the goals to be reached.

The skill component of the ADOPT appeared to have some validity in predicting real-world performance. Measures of safe driving practice, however reliable they may have been, did not appear to relate to any aspect of real-world driving performance.

Results for Individual Behaviors

Any improvement in the reliabilty and, if possible, the validity of the ADOPT required some refinement in the test measures. In order to identify what measures needed improvement and what course that improvement should take, results obtained in each of the ADOPT behaviors were analyzed.

The results of the analysis are detailed in Table 8 on the following page. All of the factors that make up the column headings represent critical aspects of the behavior measures, and all must be considered simultaneously in assessing each behavior. That is why they are displayed in a single table.

The column headings in the table are explained as follows:

Checks--The total number of times the performance was observed and scored across the three test routes. Dividing the number by 3 will give the average number of checks per route.

Scores--The number of applicants having the score; that is, the number of applicants for whom at least one observation of the behavior occurred. A behavior that cannot be scored for every driver tends to detract from the uniformity of the test.

Level of Performance--The proportion of times the behavior was performed correctly, averaged across all applicants. Behaviors that are performed correctly almost all the time by all applicants do not differentiate between good and poor drivers.

Variability of Performance--The variability of performance around the mean. Behaviors that show very little variability do not differentiate among subjects.

Examiner Reliability--The intercorrelation of scores on the behavior given by front and rear seat examiners. A low correlation indicates a rather subjective measure of behavior.

Internal Consistency--The correlation between any single measure of behavior and score on the total test or a subtest. Low correlations do not mean the behavior is unimportant, but cast some doubt upon the ability of the behavior to be combined with other behaviors in generating an estimate of proficiency.

TABLE 8

RESULTS OF ANALYSIS FOR INDIVIDUAL BEHAVIORS

BEHAVIOR	CHECKS	SCORES	PERFORMANCE LEVEL	PERFORMANCE VAR.	EXAM REL.	INTERNAL CONSISTENCY TOTAL	INTERNAL CONSISTENCY SKILL	INTERNAL CONSISTENCY PRACT	SCORED VALDITY SKILL	SCORED VALDITY PRACT	RATED VALIDITY SKILL	RATED VALIDITY PRACT
SKILLS												
BA-Brake Application	23	150	0.86	0.20	0.48	0.48	0.57	0.24	-.24	-.25	-.09	-.13
CS-Controlled Stop	6	149	0.43	0.44	0.71	0.36	0.30	0.30	0.14	0.05	-.07	-.08
RS-Rapid Stop	3	135	0.36	0.48	0.75	0.19	0.24	0.07	-.19	-.10	-.25	-.15
LS-Lane Keeping Straight	12	150	0.91	0.19	0.34	0.38	0.41	0.18	-.39	-.34	-.25	-.33
LC-Lane Keeping Curve	10	150	0.64	0.37	0.43	0.36	0.35	0.25	-.11	-.10	-.16	-.09
LI-Lane Keeping Intersect	26	150	0.70	0.24	0.60	0.53	0.60	0.25	-.30	-.35	-.04	-.23
GS-Gap Selection	14	141	0.65	0.38	0.44	0.14	0.23	-.03	-.25	-.25	-.19	-.18
SC-Speed Curve	35	150	0.84	0.18	0.31	0.59	0.78	0.18	-.21	0.14	-.02	0.11
HT-Handling Time	3	149	0.72	0.45	0.82	0.34	0.27	0.28	-.22	-.23	-.16	-.25
HD-Handling Direction	3	148	0.72	0.45	0.74	0.25	0.26	0.15	-.09	-.21	-.09	-.23
HS-Handling Strikes	3	149	0.62	0.49	0.70	0.10	0.12	0.05	0.06	-.02	0.10	0.05
HO-Handling Observing	3	149	0.82	0.39	0.45	0.33	0.34	0.22	-.21	0.00	-.09	-.04
PRACTICES												
WG-Wheel Grip	23	150	0.91	0.24	0.75	0.33	0.10	0.48	0.04	0.04	0.01	-.03
SL-Speed Limit	4	149	0.95	0.20	0.45	0.26	0.14	0.32	0.12	0.19	0.11	0.11
SM-Speed Maintenance	6	149	0.91	0.20	0.45	0.08	0.04	0.10	-.07	0.11	-.37	-.23
SI-Speed Intersection	7	150	0.38	0.36	0.36	0.38	0.11	0.55	0.03	-.13	0.03	0.06
OB-Observing Behind	13	148	0.42	0.34	0.43	0.41	0.30	0.40	-.11	0.00	-.06	-.15
OR-Observing Quarter	9	150	0.64	0.43	0.88	0.32	0.10	0.44	0.21	0.18	0.11	0.05
OS-Observing Side	16	150	0.52	0.26	0.75	0.42	0.27	0.45	-.22	-.20	-.10	-.22
CL-Communicate Lane	11	150	0.82	0.25	0.64	0.26	0.05	0.37	0.05	0.08	-.19	-.13
CT-Communicate Turn	26	150	0.97	0.07	0.27	0.26	0.12	0.33	0.21	0.11	0.02	0.09
DR-Distance Restriction	10	150	0.75	0.30	0.61	0.45	0.27	0.51	0.27	0.17	0.05	0.15
DA-Distance Adjacent	9	150	0.89	0.22	0.40	0.30	0.19	0.30	0.10	0.08	0.17	0.09
DO-Distance Oncoming	14	149	0.96	0.11	0.13	0.37	0.23	0.39	0.13	0.13	-.14	-.02
DF-Distance Following	9	78	0.83	0.36	0.43	0.08	-.01	0.14	0.46	0.04	0.65	0.18
TOTALS												
SKILLS TOTAL	142	150	0.76	0.14	0.61	0.85	1.00	0.40	-.41	-.30	-.18	-.23
PRACTICES TOTAL	156	150	0.78	0.10	0.79	0.82	0.40	1.00	0.17	0.10	-.01	-.04
URBAN TOTAL	92	150	0.81	0.11	0.72	0.81	0.62	0.73	-.08	0.00	-.07	-.11
SUBURBAN TOTAL	142	150	0.78	0.11	0.74	0.87	0.77	0.70	-.27	-.24	-.08	-.08
RURAL TOTAL	56	100	0.69	0.13	0.65	0.81	0.71	0.57	-.05	-.04	-.17	-.25
TOTAL	298	150	0.77	0.09	0.76	1.00	0.85	0.82	-.25	-.20	-.16	-.23

Scored Validity--The correlation of the behavior with scores obtained from analysis of videotaped real-world performance. Since the ADOPT is scored in terms of correct performance and the videotapes were scored in terms of errors, a negative relationship would be expected.

Rated Validity--The correlation of behavior with ratings of videotaped real-world behavior by the expert panel. Because of the nature of the rating process, a negative correlation would be anticipated.

Results for each of the factors listed in the table were obtained separately for front and rear seat examiners, with the obvious exception of the correlation between the two examiners. Because the examiner in administering a real road test would always occupy the front seat, it is the front seat examiner's results that are displayed in the table. Each of the test behaviors will be assessed relative to these factors in the following paragraphs.

Brake Application

Brake application evidenced only marginal variability (.20) and examiner reliability (.48). However, it showed a high correlation with the overall skill component of the ADOPT as well as validity coefficients that were in the correct direction.

Scoring brake application required a judgment on the part of the examiner as to smoothness or abruptness of deceleration. While use of an accelerometer in training produced fairly high levels of agreement by the end of the training period, it is clear that the agreement was not sustained. Because of the promising validity relationships and internal consistency, this behavior was retained in the ADOPT but with the recognized need to improve interexaminer agreement through additional training.

Controlled Stop

This behavior resulted in fairly high variability (.44) and high examiner reliabilty (.71). However, this is countered by a low level of internal consistency and marginal validity coefficients.

The low mean performance probably indicates the failure on the part of applicants to understand what they were supposed to do rather than any limitation in their braking skill. Most of the locations at which the behavior was called for were characterized by a lack of an identifiable stopping point other than the extension of the sidewalk line into the street.

Even where there was a painted stop line, there was no reason to expect a driver to stop within any specified distance of it, provided the line was not exceeded. Controlled Stop could have measured braking skill if drivers were specifically told to bring the vehicle to a stop at a delineated point. However, such a contrived maneuver becomes an exercise rather than an element of ordinary street operation. For this reason, Controlled Stop was eliminated from the ADOPT.

Rapid Stop

The availability of very clear scoring criteria yielded a high level of examiner agreement (.75). The behavior also correlated with the two validity measures. However, a low mean score (.36) indicated that applicants were having trouble with this part of the test. The difficulty seemed to lie in the braking distances allowed under the scoring standards; they were simply too short for most applicants. It was decided that while the distances should be increased somewhat, it was desirable that they not be too accommodating. Otherwise, they would provide no incentive to drivers to practice and develop their emergency braking skills before taking the test.

Lanekeeping: Straight

This behavior produced a very high level of performance (.91) and low variability (.19). Applicants did not have too much difficulty keeping the vehicle in the lane where it was being operated in a straight path. However, the errors that were made appeared to be highly diagnostic of poor real-world performance since the validity coefficients were among the highest obtained. The behavior correlation with the skill component of the ADOPT was also quite high (.41). Lanekeeping: Straight appeared to be appropriate for inclusion in the ADOPT just as it was.

The low interexaminer agreement (.34) probably does not reflect as badly on the reliability of the measure as it would seem. Early in test development it was found that a front seat observer could judge lane position quite accurately while a back seat observer had a great deal of difficulty.

Lanekeeping: Curve

The relatively low performance (.64) and high variability (.37) of this behavior indicated that it is more difficult for novices to stay within the lane on a curve than when going in a straight line. This difficulty results in a high degree of variability. While the validity coefficients are much lower than those for lanekeeping in a straight line, they are at least in the right direction. Given these relationships, and the fact that examiner reliability was probably higher than the correlation of .43 suggests, the behavior appeared to merit retention in the ADOPT.

Lanekeeping: Turn

Results for this behavior were similar to those of the other two lanekeeping behaviors. While the variability was only moderate, examiner reliability, internal consistency, and validity coefficients were among the highest. This may be due in part to the number of times this behavior was checked--over 6 times per route. Regardless of the reason, it warranted inclusion in the ADOPT.

Gap Selection

The biggest problem with this behavior was the inability to assure that an applicant would face a Gap Selection task some time during the road test. Nine applicants were not presented with one at any time in their 20 minutes of driving. The small number of observations may be responsible for the low examiner reliability and internal consistency correlations.

Errors in Gap Selection almost always involve being overly cautious, i.e., passing up a safe gap. The videotapes of nontest behavior revealed a tendency for excessive caution to be associated with a lack of vehicle handling skill. This may help account for the correlation between Gap Selection errors and the validity measures. Whatever the reasons for these correlations, they were deemed high enough to warrant continuing this behavior in the ADOPT through the field testing phase.

Selecting Speed: Curve

Results from this measure resembled closely those obtained from Brake Application. The two are companions in that they require a judgment of acceleration by the examiner. As in the case of Brake Application, these judgments needed to be improved. However, the high level of internal consistency and the fact that validity coefficients were in the right direction warranted continuation of this measure. This behavior was checked at both curves and turns (intersections). To avoid confusion in preparing test forms, it became two checks: Selecting Speed: Curve and Selecting Speed: Turn.

Handling

The four measures of vehicle handling--time, direction, striking, and observing--can be discussed together since they were all obtained in the same maneuver.

Time and direction gave highly similar results: moderately high variability, high examiner reliability, low correlation with the overall skill measure (internal consistency), and validity coefficients in the right direction. This similarity of results is not surprising in view of the obvious relation between the number of direction changes that are made and the time it takes to complete the parallel parking maneuver. The correlation between the two measures was .66. While this close relationship might have allowed one of the measures to be eliminated in the final ADOPT, it seemed prudent to continue collecting both measures in the field test.

Striking a traffic cone or barricade during the parallel parking maneuver resulted in fairly high variability and examiner reliability. However, this variable was essentially uncorrelated with other test measures or validity measures.

Observing yielded slightly higher coefficients of internal consistency and validity than did striking, but very low examiner reliability. The examiners did not have a difficult time determining which direction an applicant was looking. However, whether they changed their direction of observation before or after a change in vehicle direction was sometimes difficult to determine.

While striking and observing seem only marginal as measures of handling skill, they were retained in the ADOPT through the field test because the parallel parking maneuver had to be performed anyway and these two measures did not conflict with the measurement of time or direction change.

No measure of position was included in the pilot test owing to the absence of a clearly defined "curb" area where the parallel parking maneuver was performed at one of the two test sites. This problem was not discovered in time to permit the collection of position data during the pilot test. However, the position criterion was retained to permit the collection of data during the field test.

Wheel Grip

A high mean score indicated that most applicants kept both hands on the wheel throughout the test. Failure to do so correlated moderately highly with other vehicle handling practices but showed no correlation with validity measures. While this observation is easy to make, it added nothing to the ADOPT and therefore was dropped.

Selecting Speed: Limit

Only 5% of the applicants ever exceeded the speed limit in the zones where speed was observed. The number of incidences of speeding might have been increased with a greater number of checks. However, even where speeding was observed, it produced low interexaminer reliablity, internal consistency, and validity coefficents.

The low examiner reliability resulted primarily from difficulty in observing the speedometer from the passenger's seat. In many vehicles, the speedometer is recessed and not visible from the passenger's side. Even when it can be observed, parallax makes it difficult to read from an angle. While observing compliance with speed limit was not informative enough to warrant remaining a check by itself, it was added to Maintaining Speed (below) to form a single check.

Maintaining Speed

Very few drivers dropped more than 10 miles under the speed limit at locations where this behavior was checked. Difficulty in reading the speedometer again produced low interexaminer agreement. However, correlations with rated validity were somewhat promising. Moreover, there is a direct connection between skill and maintaining speed in that drivers who lack vehicle handling skill often attempt to compensate for it by driving slowly. Such compensation is healthy in ordinary driving. It is less so, however, when it covers up skill deficiencies during a license test.

In giving a test, it is reasonable to require examinees to carry out the test in the prescribed manner. Likewise, it is reasonable to require applicants taking a road test to operate at speeds that will require a demonstration of vehicle handling skills. Examiners cannot force applicants to operate at the speed limit. However, a penalty for operating at excessively low speed helps encourage maintaining an adequate speed for testing. For this reason, Maintaining Speed was retained in the ADOPT. However, it was revised to include Selecting Speed: Limit. As revised, it checked to see if the applicant's speed was no greater than 10 mph under or 5 mph over the posted limit.

In the pilot test, compliance with speed limits was checked only on the open road. However, many applicants were found to drive too fast or too slow in negotiating curves and when making right and left turns. To allow speed to be checked in these maneuvers, the check was divided into Maintaining Speed: Straight, Maintaining Speed: Curve, and Maintaining Speed: Turn.

Speed: Intersection

The low mean for this behavior (.38) does not necessarily mean that applicants were behaving unsafely. The behavior called for was a reduction in speed when approaching a blind intersection. Many of the applicants who failed to do this were driving slowly already. The low interexaminer agreement reflects some indecision on the part of examiners as to whether such behavior should be called "correct," "incorrect," or "not applicable."

Because most license applicants tend to proceed very cautiously, the need for a speed reduction is likely to be rare. For this reason, the behavior was dropped from the ADOPT.

Observing: Behind

The mean, variability, internal consistency, and validity of this behavior were all marginally acceptable. The biggest deficiency was the low interexaminer agreement. Checking this behavior required the examiner to observe the applicant's eye movements, through a mirror, at a time when other checks were also being made. Examiners often simply missed making the check. Another source of disagreement was deciding when the applicant could make the observation and still have it scored as correct. If the driver had looked behind 3 or 4 seconds ago and found no vehicle there, was it necessary to look again at the time the check was called for?

Because the behavior is important, and frequently overlooked even on a license test, it was kept in the ADOPT. However, it was recognized that examiner reliability would have to be improved by reducing the number of competing checks and sharpening scoring criteria.

Observing: Rear Quarter

Making an over-the-shoulder check before changing position on the road produced a high degree of variability, examiner reliablity, and correlation with other measures of safe practice. However, validity coefficients were low and in the wrong direction. Because of the importance of the behavior, and the fact that it does not conflict greatly with other checks, it was retained.

Observing: Side

Taking a look to the side when approaching a blind or uncontrolled intersection evidenced acceptable variablity, interexaminer reliability, and correlation with other practices. Moreover, correlations with validity measures were all in the right direction. For these reasons, the behavior was retained.

Communicating: Lane Change

This behavior evidenced marginal but acceptable levels of variability, examiner reliability, and correlation with other measures. It also yielded validity coefficients that were in the right direction. Despite its marginal value as a measure of safe driving performance, this behavior was retained for its potential as an indirect measure of skill.

Communicating: Turn

If there was one thing applicants remembered to do, it was to signal turns. Only 3% of the turns were overlooked. With very little variation in performance, there was very little correlation with other variables. Since it does not contribute to the effectiveness of the ADOPT, Communicating: Turn was dropped from further testing.

Restricted Travel

This behavior produced marginal but acceptable levels of variability, examiner reliability, and internal consistency. Validity coefficients were, if anything, in the wrong direction.

While this behavior is not particularly indicative of drivers' safety of operation in the real world, there are two reasons why it warranted retention. First, any violation of travel restrictions, such as crossing a painted gore, is illegal and potentially dangerous enough to demand recognition in test scoring. Secondly, in order to permit assessement of vehicle handling when turning a corner and in order to set up certain lane changes, it is necessary that applicants turn into the correct travel lane. Checking compliance with travel restrictions provides an incentive to perform correctly and a penalty for not doing so. The role of this check is very similar to that of Maintaining Speed in forcing drivers to perform in a way that will allow other behaviors to be assessed.

Distance: Adjacent, Oncoming

These two behaviors gave results that are sufficiently similar to allow them to be treated as one. A high level of performance was accompanied by very little variability, which in turn suppressed examiner reliability, internal consistency, and validity coefficients. The principal problem with these behaviors is that applicants who are in the right lane position to begin with don't have to do anything in order to be scored as "correct." The problem is similar to that described in the case of Speed: Intersection, except in that case drivers who were doing the right thing beforehand were scored as "incorrect."

While it is important to maintain adequate distance from adjacent and oncoming vehicles, it did not appear possible to assess the practice through an on-road test, and the behaviors were dropped.

Distance: Following

Almost half of the 150 applicants failed to encounter a following situation in the zones where a check was called for. This result is similar to that obtained in other attempts to assess following distance in a road test. With so few observations, the results obtained were highly unstable and not much confidence can be placed in the figures that are given. Therefore, this behavior was dropped from the ADOPT.

Selection of Behaviors

Based upon results just described, the following behaviors were selected for inclusion in the ADOPT for further development and field testing:

- Brake Application
- Rapid Stop
- Lanekeeping:
  - Straight
  - Curve
  - Turn
- Gap Selection
- Selecting Speed:
  - Curve
  - Turn
- Handling:
  - Time
  - Direction
  - Strikes
  - Observing
  - Position
- Maintaining Speed:
  - Straight
  - Curve
  - Turn
- Observing:
  - Behind
  - Rear Quarter
  - Side
- Communicating: Lane Change
- Restricted Travel

The following behaviors were eliminated from further study:

- Controlled Stop
- Selecting Speed:
  - Limit
  - Intersection
- Communicating: Turn
- Distance:
  - Adjacent
  - Oncoming
  - Following
- Wheel Grip

FIELD TEST

Following its revision on the basis of pilot test results, the ADOPT was field tested in Oklanoma.

PURPOSE

The purpose of the field test was threefold:

- o Assessment of reliability
- o Assessment of validity
- o Comparison of reliability and validity with State road test

Reliability

A number of revisions to the ADOPT were intended to improve its reliability. The changes included (1) revising test administration methods in order to improve tne objectivity of examiner judgments and (2) eliminating those checks did not appear to be amenable to objective measurement. The same aspects of reliability that were assessed in the pilot test were reassessed in the field test. These included examiner reliability, sampling reliability, and measurement reliability.

One change was made--an alteration in the way sampling reliability was measured. During tne pilot test, samples of behavior were compared across three different environments: urban, suburban, and rural. This was done in order to determine the effect of the different environments upon reliability. Such a comparison was consistent with the objective of the pilot test, which was primarily to learn things about the components of the ADOPT that could lead to improvements. However, the objective of the field test was not so much discovery as it was to provide an assessment of the completed ADOPT. A more meaningful measure of sampling reliability was the relationship among complete routes. Therefore, in the field test, sampling reliability became a correlation between scores obtained by the same applicants driving over two different test routes.

Validity

The pilot test revealed small but significant correlations between the ADOPT and measures of real-world operating skill. No significant correlation appeared between the ADOPT and measures of real-world safe operating practice.

The measure of real-world skill employed in the pilot test was necessarily somewhat crude. It is difficult to measure a driver's vehicle handling skill from a following vehicle. In the field test, a more precise measure of skill was employed. It involved (1) following and videotaping applicants along the same route in order to eliminate the effects of route differences and (2) testing applicants in an off-street environment in which more challenging tasks could be created and applicant performance could be measured more precisely.

Comparison with State Test

Since States already administer road tests to license applicants, if the ADOPT is implemented in any State, it must be as a replacement for the test currently given. Such replacement would be warranted only if the ADOPT represented an improvement over the current test.

The ADOPT could not be compared with road tests employed in all States. However, as has been mentioned earlier, there is a great deal of similiarity among the road tests employed by different States. In contrast with the ADOPT:

- They generally employ a subjective assessment of applicant performance rather than an objective assessment against specified criteria.
- The performances checked are largely confined to those aspects of vehicle operation that are most easily measured.
- Overall test score does not attempt to take variation in traffic conditions into account (nor is such really necessary, given the subjectivity of the performance assessment).

This description of State tests in general also characterizes the test employed by Oklahoma. Therefore, the reliability and validity of the Oklahoma road test was analyzed and compared with that of the ADOPT.

METHODS

Selection of Field Test Site

Oklahoma was selected for the field test for the same reasons given earlier in this report in the description of the pilot test. In addition, it was possible to use some of the personnel who had participated in the pilot test and to take advantage of the experience gained by the Department of Public Safety during administration of the pilot test. The site selected for the field test was Tulsa, which was one of the two locations used in the pilot test. As mentioned earlier, this licensing station averaged 100 applicants per week, more than any other station in the State, and therefore permitted sampling requirements to be met in the shortest possible time.

Route Selection

For the field test, project staff selected a total of four routes. Two were used for administration of the ADOPT; two were used for the State test.

ADOPT Routes

These routes were not the same as the routes used in the pilot test. A key difference was in driving time per route; the field test routes required 10 minutes (including parallel parking), whereas the pilot test routes required 20 minutes. Total driving time was actually 20 minutes in the field test, however, since applicants drove two routes in order to permit sampling reliability to be assessed.

The two routes involved mostly low-speed driving in business and residential areas, with brief stretches where speed limits were as high as 45 mph. The routes were 3.2 and 3.0 miles long. They were chosen to permit nearly equal number of performance checks to be made.

State Test Routes

The two State test routes were located entirely in a residential area. The routes were 1.1 and 1.0 miles long, requiring 6.5 and 6 minutes of driving, respectively, including parallel parking. The routes were chosen from among routes regularly used by the DPS to administer the State road test.

Skill Test Facility

As noted earlier, an off-street Skill Test was employed to provide a more precise measure of handling skill. It also established a destination for a fixed route of travel, minimizing differences in the conditions encountered by drivers whose performance was being videotaped.

Since the site of the Skill Test determined the route that would be taken to reach it, it would have been a great advantage to be able to select the site so that the route to the site from the licensing station permitted videotaping of a wide variety of driving behaviors. It was necessary, however, to choose a site that was relatively close to the licensing station (since if it were too far away, applicants would be reluctant to participate) and that provided an area suitable for the Skill Test. The site chosen was a public school bus garage and parking lot approximately a 5-minute drive from the licensing station. The route for reaching the site provided opportunity for a sufficient number of driving behaviors.

Field Test Sample

The field test sample consisted of 300 subjects, all novice drivers applying for a license for the first time. The ADOPT and the State road test each were administered to 150 applicants. In all cases, an attempt was made to select for testing those applicants who arrived at the licensing station unaccompanied by an adult (or other authority figure). This was an effort to maximize the number of applicants whose free, nontest driving could be videotaped.

Processing of Applicants

As indicated above, those applicants arriving unaccompanied by an adult made up the subjects sample. At the licensing station, applicants were provided with a sheet of paper explaining how their road test would be scored. Applicants were instructed to read the explanation and place it under the windshield wiper of their vehicle when they were ready to take the test. Examiners tested applicants who displayed the explanation sheet on their vehicles. This continued until 300 subjects had been tested.

Road Test Administration

The examiners worked in pairs, one sitting in the passenger seat in the front of the car, the other riding in the rear center seat. They reversed positions prior to each successive test. A plausible explanation for the presence of the second examiner was given to the applicant (e.g., on-the-job training).

Applicants taking each test were scored on both of the routes making up that test. For each test, the order in which the routes were given was alternated with each successive applicant.

Scoring

Applicants were scored by both examiners. During the State test administration, examiners used only the State score sheet. During administration of the ADOPT, examiners used only the ADOPT test form, but completed the State score sheet (on the basis of first-route performance only) once the ADOPT administration had been completed. Pass/fail determination was made on the basis of the State score achieved on the first route driven.

Obtaining Additional Measures

Once the road test had been administered and scoring was complete, applicants were informed whether they had passed or failed. Those applicants who had arrived at the site unaccompanied by an adult were approached by an examiner who solicited their participation in the off-street Skill Test.

- Applicants who had <u>failed</u> the road test were told that a short, free training program was being given nearby, and that participating in the program might help them to pass the road test on their next attempt.

- Applicants who had passed the road test were told that the State was developing an off-street skill test at a nearby site, and that their participation would help the project.

Collecting Additional Data

Those who agreed to participate were given a map showing the route from the licensing station to the Skill Test site. This was done in an effort to achieve uniformity of routes among applicants, as it was assumed that most would follow the map. The field assistant then followed the volunteer surreptitiously, videotaping along the route to the Skill Test site. The procedures and equipment used in videotaping were identical to those used in the pilot test. The examiner also contacted the skill test administrator to inform her that a volunteer was en route to the site.

Videotaping

Taping began with the applicant's exit from the licensing station parking lot. Applicants had to back out of an angled parking space and proceed through the lot to enter the street. Taping continuously, the field assistant followed the applicant to the Skill Test site, regardless of the route taken to the site. When applicants reached the test site, they followed signs directing them to the test location. The signs led them through contrived maneuvers--around a parked vehicle and between barricades--to a large sign that directed them to back into a designated parking slot. Once parking was complete, taping ceased.

Up until the time they parked the car, applicants had no idea their performance was videotaped. Therefore, all aspects of their driving, even the contrived maneuvers, can be considered representative of real-world performance.

It was not always possible to keep applicants within the view of the video camera during the contrived maneuvers. Therefore, the skill test administrator, out of sight of the applicants, also scored their performance, including parking performance. Scoring noted the number of stops and direction changes and the time for completion.

Administering Skill Test

After applicants had parked, the administrator approached and explained that they would perform three maneuvers: serpentine, "T" exercise, and head-in parking. Once the maneuvers had been completed, applicants were told the test was over and were thanked for their participation. In the case of an applicant who had failed the road test, the administrator discussed each maneuver and gave tips for improving performance.

While only these three maneuvers comprised the Skill Test as far as the applicant was concerned, three additional maneuvers were scored as a part of the test: maneuvering around the barricade, the back-in parking, and backing-out.

Serpentine

The path of travel was established by a straight line, 120 feet in length, created by markers (3 feet tall) spaced 30 feet apart. Side boundaries were established by markers placed in straight lines 25 feet to the left and right of the center markers. This created an area approximately 50 feet wide and 120 feet long. The subject was positioned to the right of the first center marker. He then drove forward to the left of the next marker, to the right of the next, and so on. Measures included: time, striking the serpentine markers, exceeding boundaries, and number of stops.

"T" Exercise

The path of travel for this maneuver was shaped like the letter "T," and was delineated by markers, 3 feet tall, and painted lines. The bottom of the "T" was 8 feet wide and 12 feet deep. The top of the "T" was 9 feet wide and 26 feet long. Subjects drove forward into the "T," turned left or right along the top of the "T," stopped, backed to the opposite end, stopped, turned, and drove forward out of the "T." Measures included: time, exceeding boundaries, and direction of observation.

Head-in Parking

This maneuver used an existing standard perpendicular parking space, 8 feet wide and 12 feet long. An end space was selected to minimize participant confusion. Subjects drove forward into the space, stopped, and backed out. Measures were taken during entry to the space and included: time, exceeding boundaries, number of stops, number of direction changes, and position of vehicle when stopped.

Barricade Maneuver

Applicants were scored on whether or not they stopped in maneuvering around the barricade positioned at the entry to the Skill Test area.

Back-in Parking

The back-in parking area was a standard 8' x 12' perpendicular parking stall. A set of cones three feet apart marked the edges of the stall. Applicants were scored on the number of cones struck during the parking maneuver, whether or not they came to rest inside the delineated parking area, and the final alignment of their vehicles.

Backing Out

One additional maneuver in the Skill Test occurred after applicants were told the test had concluded. Their performance as they backed out of the parking area was surreptitiously observed and scored by the administrator. The only scoring criterion employed was exceeding the boundaries.

Scoring Skill Test

The individual exercises making up the Skill Test had widely differing variances. For example, the time to complete a maneuver, in seconds, had a much wider range of values than did number of cone strikes. Witnout some equalizing procedure, a few of the exercises could easily have dominated the total score. To keep this from happening, scores on the individual measures were all converted to standard scores having the same variance.

Most of tne scoring criteria were such that high values were associated with low skill, e.g., time to complete the maneuver, number of cone strikes. For tne remaining criteria, the direction of scoring was somewhat arbitrary, e.g., whether or not the driver's hand position was correct. These were also scored so that incorrect performance carried a high value.

Field Test Preparation

Project staff made two site visits to Oklahoma to prepare for the field test. During the first visit, DPS personnel were briefed on the nature and purpose of the field test and on the ADOPT. Administrative procedures were explained, routes were selected, and examiners for the field test were designated. The skill test administrator and the field assistant were hired.

The second site visit was made just before the field test began. At that time, all field test personnel were trained thoroughly by members of the NPSRI project staff. Field test personnel consisted of four examiners (one of whom was designated the coordinator), a field assistant, and a skill test administrator. NPSRI staff provided overall supervision.

Examiners

The examiners were divided into two teams of two examiners each. One team administered the ADOPT while the other administered the State test. Two of the four examiners had participated in the pilot test. The examiners were paired so that there was an "experienced" examiner on each team. In addition to administering the road tests and scoring the applicants, tne examiners were responsble for obtaining the volunteers for the Skill Test.

Examiners received two days of training, most of which was devoted to practice administrations. Examiners first administered the ADOPT to surrogate applicants (project staff members) and eventually to actual applicants. Practice continued until an interexaminer reliability level of .8 was reached.

While no training on the State test routes was necessary, project staff members drove over the two routes with the examiners (1) to ensure that the routes were followed exactly, (2) to determine where performance checks would be made, and (3) to ensure that performance checks were made uniformly and consistently by all examiners.

Coordinator

One of the four examiners was designated the coordinator. His responsibilities included checking test forms for completeness, collecting videotapes and test forms and mailing them to NPSRI on a weekly basis, and contacting NPSRI periodically to report on the field test status and the number of applicants tested to date. Project staff reviewed all procedures thoroughly with the coordinator during the second site visit.

Field Assistant

The field assistant was engaged locally and was not connected with the Department of Public Safety. Her responsibility was videotaping the volunteers as they drove from the licensing station to the site of the Skill Test. The field assistant was trained by project staff and performed several practice tapings to become familiar with procedures. The practice tapes were reviewed to make sure that required data were being captured and all equipment was functioning properly.

Skill Test Administrator

Like the field assistant, the skill test administrator was hired locally and had no connection with the Department of Public Safety. She was trained in all procedures to be followed in administering the skill test to the volunteers and in scoring their performance. Practice was extensive, covering nearly three days. Initially, project staff served as "applicants." Eventually, as the administrator gained experience in procedures, actual applicants participated in the skill test. This had the advantage of ensuring the method of obtaining volunteers was feasible and would be successful.

Analysis of Videotapes

As in the pilot test, the videotape records of drivers followed on their departure from the licensing station were subjectively rated and objectively scored. The objective scoring process was identical to that used in the pilot test. However, a variation of the rating process was necessary. Over ten hours of videotape were accumulated on a total of 142 applicants. If it required as much time to carry out the rating process in the field test as it had in the pilot test, it would have taken the panel some four days to rate the performance of the applicants. The original panelists were not available for this period of time, nor could the project afford such an effort. Moreover, because of the high interrater reliability obtained in the pilot test, a large number of panelists seemed unnecessary. Therefore, two staff members performed the rating, using the same process employed during the pilot test.

RESULTS

Results from the field test like those from the pilot test were analyzed to assess examiner reliability, sampling reliability, measurement reliability, and external validity. The same classification of behaviors into Skills and Practices was maintained.

The primary difference between the pilot test and the field test was the comparison of the ADOPT with the Oklahoma State examination. Each of the analyses is therefore presented for both the ADOPT and the Oklahoma test (OKLA).

Because the objective of the field test was to evaluate the ADOPT as a whole, the analysis of data focuses upon overall test scores. While there is frequent mention of individual behavior categories, there is no need for the detailed analyses of each behavior category that were performed in the pilot test. Those analyses were performed in order to aid in selection of individual behaviors.

Examiner Reliability

Examiner reliability was again analyzed in terms of both constant and variable error.

Constant Error

Constant error was examined by comparing mean scores given by the two examiners assigned to each test and each route. The results appear in Table 9 below.

TABLE 9
MEAN SCORE BY EXAMINER POSITION

	ADOPT		OKLA	
Measure	Front	Rear	Front	Rear
Total	.77	.77	.95	.95
Skills	.82	.82	.22	.22
Practices	.58	.56	.37	.37

Since the ADOPT score is expressed in terms of percentages, the total is the average of Skills and Practices. Oklahoma test is scored in terms of points and the two components are additive. The results show that the two examiners administering each test averaged exactly the same total score and almost the same score on each component.

Within the individual behavior categories on the ADOPT, the differences between examiners are also very small. In no category did the means differ by more than 5% and in most cases it was 3% or less.

Variable Error

Variable error is the tendency for different examiners to give different scores to the same applicants. Unlike constant error, it cannot be corrected for.

As in the pilot test, variable error was obtained by correlating scores of front and rear examiners. The results appear in Table 10 below.

TABLE 10
INTERCORRELATION OF EXAMINERS SCORES

	ADOPT			OKLA		
Measure	Rt. 1	Rt. 2	Total	Rt. 1	Rt. 2	Total
Total	.81	.87	.84	.85	.84	.84
Skills	.80	.86	.83	.82	.78	.80
Practices	.75	.72	.74	.63	.51	.57

Examiner reliability exceeds the .8 standard that was set in advance. While the improvement over the pilot test score appears relatively modest, it is important to remember that the length of the route has been cut in half--from twenty minutes to ten minutes.

In the field test, the examiner's judgment of Skills proved more reliable than judgment of Practices. This is a reversal of the findings in the pilot test, where examiner reliability for Practices was substantially higher than that for Skills. The change probably reflects (1) improvements in the criteria for scoring skills and (2) the fact that skills are much more substantially sampled than practices in the final version of the ADOPT.

The examiner reliability of the OKLA test was identical to that of the ADOPT--.84. Like the ADOPT, the OKLA gave more reliable judgments of Skills than Practices. What is surprising is that the reliability of the overall OKLA is identical to that of the ADOPT while both components of the OKLA showed less reliability--particularly Practices.

Examiner reliabilities for individual measures were considerably higher than those obtained in the pilot test. Most exceeded .5 and some were over .8. The high reliability of individual behavior categories is probably a joint function of (1) improving the objectivity of scoring criteria, and (2) eliminating the behaviors with extremely low reliability.

Summary

In summary, it appears that both the ADOPT and the OKLA lead to highly similar scores when administered by different examiners. A reliability coefficient of .84 for total score is extremely high for a ten-minute road test.

Sampling Reliability

Sampling reliability was determined by comparing scores on the two routes over which each applicant drove.

Constant Error

The constant error in the case of sampling reliability is the tendency for one route to give scores that are consistently different from another route. It was analyzed by comparing mean scores on both routes, regardless of examiner. Results appear in Table 11 below.

TABLE 11
MEAN SCORES BY ROUTE

Measure	ADOPT Rt. 1	ADOPT Rt. 2	OKLA Rt. 1	OKLA Rt. 2
Total	.78	.78	.59	.59
Skills	.83	.81	.22	.22
Practices	.57	.57	.37	.37

ADOPT scores over the two routes gave very similar mean scores. Route 2 may have been slightly more difficult from a skill viewpoint, but the difference is negligible. The OKLA gave identical scores on both routes.

Variable Error

Variable error in the case of sampling reliability is the tendency for various samples to give different results for different applicants. It was assessed by intercorrelating scores achieved by applicants over the two routes they drove. Results appear in Table 12 below.

TABLE 12
INTERCORRELATION OF SCORES ACROSS ROUTES

	ADOPT			OKLA		
Measure	Front	Back	Total	Front	Back	Total
Total	.76	.76	.76	.82	.83	.82
Skills	.76	.76	.76	.71	.77	.74
Practices	.51	.45	.48	.70	.66	.68

The sampling reliability of the ADOPT as a whole is quite high, far exceeding the .6 that was established as a target. The improvement in sampling reliability over the pilot test can be explained, at least in part, by the fact that this sample consisted of two routes designed to be similar rather than three rather different environments (i.e., urban, suburban, and rural). The reliability of Skills is again higher than that of Practices. On the ADOPT, the reliability of the Total score was equal to that of Skills. Sampling reliability of the OKLA test is also high, slightly higher than that of the ADOPT.

Summary

Both the ADOPT and OKLA produced similar results when administered over different routes. High sampling reliability for the OKLA is not surprising given the somewhat subjective scoring system. An examiner having assigned a score on the first route is likely to be somewhat influenced by that score when scoring the performance on the second route. Since scores on the ADOPT are determined much more by specific situations occurring along a route, a high correlation is to a greater degree an indication of the functional similarity of the routes themselves.

Measurement Reliability

Measurement reliability refers to the extent to which applicants taking the test from one examiner on one route achieve similar scores from another examiner on another route. It was estimated by obtaining the correlation between all six possible pairs of examiners and routes, i.e., Examiner 1, Route 1 versus Examiner 2, Route 2, and so on. The results were as follows:

TABLE 13
MEASUREMENT RELIABILITY

Measure	ADOPT	OKLA
Total	.76	.77
Skill	.75	.70
Practices	.46	.46

The measurement reliability of the ADOPT is substantially higher than the target of .5 and almost twice that obtained in the Pilot Test. The increase is the result of the improvements in examiner and sampling reliability that have been described. The measurement reliability for the OKLA and the ADOPT are almost identical.

On both the ADOPT and OKLA, applicants can be reasonably assured that they will obtain the same score no matter which route they travel and which examiner administers the test. This conclusion does, however, have its limitations. The pairs of examiners and routes involved in each test represented only one license station. Had the applicant gone to a another license station to take the test from a different examiner over a different route, the correlation might not have been as high. On the other hand, both the examiners and the routes appeared to be reasonably representative of examiners and routes elsewhere. Differences are not likely to produce a large reduction in measurement reliability.

Intercorrelation of ADOPT and OKLA

All applicants administered the ADOPT were also scored in the OKLA in order to determine their eligibility for a license. The intercorrelation of scores of the two measures is shown below:

TABLE 14
INTERCORRELATION OF ADOPT AND OKLA

Measure	Correlation
Total	.52
Skills	.45
Practices	.27

While the two tests did not give the same results, there is a moderate correlation between them. It is hard to say how much of this correlation is due to similarities between the tests and how much is due to the fact that the same examiner scored both of them. The OKLA score was determined by the front seat examiner after administering the ADOPT. In scoring the OKLA, the examiner may well have been influenced by the scoring of the ADOPT.

Correlation with Skill Test

The off-street Skill Test was intended to provide a more accurate measure of operating skill than could be observed by following drivers on the open road. Given the fact that highways are designed to minimize the demands placed on drivers' skills, it is very difficult to make an accurate assessment of skill in normal driving.

Skill Test Sample

Some 74 of the 300 applicants completed the off-street Skill Test. The remaining applicants fell into the following categories:

- Arrived with an adult and were not asked to take the test.
- Would not agree to take the test.
- Volunteered to take it but failed to show up.

Of the 74 applicants taking the Skill Test, 34 had taken the ADOPT and 40 had taken the OKLA. The representativeness of the Skill Test samples can be seen in the comparison of Total and Skill Test samples in Table 15 below.

TABLE 15
MEAN SCORES OF TOTAL AND SKILL TEST SAMPLES

	ADOPT		OKLA	
Measure	Total	Skill	Total	Skill
Total	.77	.77	.59	.60
Skills	.82	.82	.22	.22
Practices	.57	.60	.37	.38

The variances showed the same high similarity as did the means. The Skill Test sample can therefore be considered representative of the Total sample when it comes to measured road test performance.

Skill Test Correlations

Of primary concern is the set of correlations between the road test scores and the scores received by applicants taking the Skill Test. These correlations are presented in Table 16 below.

TABLE 16
CORRELATION OF ADOPT AND OKLA
ROAD TESTS WITH THE SKILL TEST

Measure	ADOPT	OKLA
Total	.56	-.01
Skill	.48	.05
Practices	.43	-.11

These results show a moderately strong association between the ADOPT scores and scores on the Skill Test. Scores on the OKLA, on the other hand, showed no relation to the Skill Test. The results seem to indicate that the ADOPT is a measure of skill while the OKLA is not.

It is important to bear in mind that the Skill Test was completely independent of the ADOPT and the OKLA. There was no communication whatever between the license test examiners and the Skill Test administrator. Therefore, Skill Test scores are totally uncontaminated by results obtained from the two road tests.

It is surprising to find that the Practices subtest correlates with the Skill Test almost as strongly as does the Skills subtest. This finding conflicts with the very low correlation between the Skills and Practices subtests on the ADOPT. One reasonable explanation is that the Skill Test, being a more valid measure of operating skill than the Skills subtest of the ADOPT, is more sensitive to the skill component of safe operating practices. The correlation between Practices and Skills found in the testing of motorcycle license applicants mentioned earlier involved an off-street skill measure and an on-street measure of practices.

The low correlation of the OKLA with the Skill Test justifies reconsideration of the reliability coefficients obtained in administration of the OKLA. The basis for the agreement between examiners and the consistency across routes was not, evidently, the skill of the applicants. This does not mean there is anything false or misleading about the reliability coefficients. Examiners could be reporting accurately what they saw, and what they saw may have been highly consistent across routes. However, what they were observing was not, apparently, highly related to the applicants' operating skill.

In a measure that is somewhat subjective, there is always an opportunity for other factors to creep into the assessment process--the applicant's cooperativeness, age, or apparent nervousness. It is the influence of such factors that the ADOPT attempts to neutralize by a highly objective scoring system.

Correlation with On-Street Performance

Neither the ADOPT nor OKLA test measures evidenced a significant correlation with videotaped records of on-street performances. The lack of significant correlations applied to both scored and rated performance and to both vehicle operating skills and safe operating practices.

Failure to find a relationship between test scores and safe operating practices employed by applicants after they left the licensing station was consonant with both expectation and results of the pilot test. The use of safe operating practices by applicants when they are being tested simply bears no relation to their behavior when they are on their own.

Relationship between test measures of both Skills and Practices on the one hand and videotaped records of driving skill on the other was anticipated on the basis of the pilot test results. While failure to find such a relationship was unexpected, it is easy to explain. It is very difficult to determine a driver's skill simply through observation of routine driving, particularly when the only element of skill that can be readily observed is lanekeeping. Finding a correlation during the pilot test was either due to chance variation in the applicant population or the conditions of measurement.

Results for Individual Behaviors

Statistics for individual behaviors appear in Table 17 on the following page. Results from both examiners and each route are pooled in arriving at the statistics shown in the table.

The contents of the table are similar to those of the table provided for analysis of individual behaviors in the pilot test. Since the overall ADOPT failed to show significant correlation with data obtained from the videotape records, correlations for individual behaviors are not presented as they were in the pilot test. On the other hand, sampling reliabilities and correlations with the Skill Test have been added to the table for all behaviors. Negative correlations between individual behaviors and the Skill Test scores are to be expected based upon the scoring system.

As noted earlier, individual behaviors were not analyzed for the purpose of determining whether they would be included in or excluded from the ADOPT. No changes have been made in the ADOPT from the form in which it was field tested. Examination of individual behaviors indicates that all are contributing to some extent to the reliability and validity of the measure. Some evidence low levels of examiner or sampling reliability, others low levels of internal consistency, and others low or nonexistent correlations with the Skill Test. However, no behavior falls short on more than one of these statistics.

It is probably unwise to attribute a great deal of significance to the specific correlations appearing in the table. While each correlation provides an estimate of a relationship for an individual behavior, each also represents a chance fluctuation around the relationship for the test as a whole, shown at the bottom of the table.

TABLE 17

STATISTICS FOR INDIVIDUAL BEHAVIORS

CATEGORY	CHECKS	PERFORMANCE MEAN	SD	RELIABILITY EXAM	SAMP	INTERNAL CONSISTENCY SKILL	PRACT	TOTAL	SKILL TEST
HT-Handling Time	2	0.94	0.24	0.39	0.12	0.22	0.08	0.22	-.28
HD-Handling Direction	2	0.94	0.24	0.66	0.28	0.24	0.15	0.26	-.58
HS-Handling Strikes	2	0.74	0.44	0.86	0.18	0.25	0.00	0.21	-.08
HO-Handling Observing	2	0.72	0.45	0.45	0.33	0.36	0.14	0.36	-.34
HP-Handling Position	3	0.25	0.43	0.88	0.45	0.31	0.15	0.32	-.19
BA-Brake Application	6	0.95	0.15	0.67	0.43	0.40	0.04	0.37	-.59
RS-Rapid Stop	2	0.34	0.47	0.89	0.47	0.33	0.07	0.32	-.35
LS-Lanekeeping Straight	8	0.95	0.16	0.56	0.44	0.43	0.01	0.39	0.06
LC-Lanekeeping Curve	2	0.77	0.42	0.67	0.17	0.36	0.09	0.36	-.12
LT-Lanekeeping Turn	8	0.68	0.35	0.75	0.72	0.50	0.10	0.47	0.04
GS-Gap Selection	5	0.57	0.46	0.82	0.23	0.33	0.20	0.35	-.61
MS-Maint Speed Straight	9	0.87	0.22	0.80	0.61	0.47	0.10	0.45	-.24
MC-Maint Speed Curve	4	0.90	0.23	0.46	0.26	0.46	0.05	0.43	-.17
MT-Maint Speed Turn	9	0.67	0.34	0.81	0.79	0.62	0.03	0.56	-.27
SC-Select Speed Curve	5	0.92	0.20	0.71	0.23	0.28	0.05	0.27	-.03
ST-Select Speed Turn	9	0.99	0.07	0.45	0.25	0.25	0.01	0.22	-.11
OB-Observing Behind	5	0.17	0.30	0.47	0.31	-.10	0.51	0.09	-.13
OR-Observing Rear	2	0.57	0.50	0.84	0.65	0.30	0.46	0.42	-.45
OS-Observing Side	5	0.56	0.38	0.79	0.32	0.08	0.61	0.29	0.08
CL-Communic Lane Change	2	0.93	0.25	0.63	0.24	0.11	0.32	0.21	-.76
RT-Restricted Travel	5	0.85	0.26	0.79	0.29	0.25	0.41	0.34	-.43
SKILLS TOTAL	78	0.82	0.11	0.83	0.76	1.00	0.17	0.94	-.48
PRACTICES TOTAL	19	0.57	0.17	0.74	0.48	0.17	1.00	0.49	-.43
TOTAL	97	0.77	0.10	0.84	0.76	0.94	0.49	1.00	-.56
SKILL TEST	62	-0.48	0.43	1.00	1.00	-.48	-.43	-.56	1.00

DISCUSSION AND CONCLUSIONS

DISCUSSION

The results of the field test showed the ADOPT to be a reliable measure of the on-road performance of license applicants. The scores obtained by applicants on the ADOPT were the same regardless of who administered the test or over what route it was administered. While the number of examiners and the number of routes were strictly limited, they were not atypical in any way. It is reasonable to believe that approximately the same levels of reliability would be realized by other examiners administering the test over other routes.

What the ADOPT Measures

With the reliability of the ADOPT established, the most important question is "What is the ADOPT a reliable measure of?"

The original source of the behaviors making up the ADOPT is a set of behaviors identified as critical to safety by a group of highway safety specialists (McKnight and Adams, 1970). The behaviors fall into two categories: those requiring the use of perceptual and perceptual motor skills, and those that combine day-to-day driving practices. The two categories of behavior may be viewed as representing, respectively, what drivers "can do" and what they "will do." Throughout development of the ADOPT, the number of behaviors was reduced as individual behaviors were selected out for lack of judged criticality to safety, ability to be measured accurately, or acceptability to licensing officials. This process did not alter the scope of the test with respect to the two basic categories of behavior.

During the pilot test, it became apparent that there was no relationship between applicants' use of safe operating practices during the test and their employment of the practices after the test was over. However, both practices and skills as measured on the ADOPT appeared to correlate with the skills observed after the test. The correlations were small, but there is little in normal driving that calls upon drivers to evidence their skill level. And only a small portion of the skill that is evidenced can be observed from outside the vehicle being driven. Use of a more precise measure of perceptual and manipulative skill in the field test resulted in more substantial correlations with the ADOPT. Again, both practices and skills as tested in the ADOPT correlated with observed skill. The results were very similar to those obtained in evaluating a road test for motorcycle operators (McPherson, McKnight, and Knipper, 1978.)

Regardless of the objective toward which it was developed, the ADOPT has emerged as a measure of driver skill. Among those behaviors in the "skill" category, such as smoothness of brake application or ability to keep the vehicle within lane, it is primarily perceptual and manipulative skill that determines how well the behavior is performed. For behaviors in the "practices" category, the relationship is less direct. There is no clear reason why drivers who signal turns or use their mirrors should be more skillful than drivers who don't. One hypothesis is that drivers who lack skill become so preoccupied with merely handling the vehicle properly that they forget to employ those safe operating practices that they know are required. More skillful drivers are better able to share attention between control of the vehicle and use of safe operating practices. While there are methods for testing this hypothesis, they are expensive to employ and were not a part of the ADOPT's development. At present, the hypothesis must be accepted or rejected on plausibility alone. Whatever the explanation, it appears that the ADOPT as a whole, as well as its major components, provides a measure of driver skill.

Comparison with Typical Road Test

The regular Oklahoma road test exhibited the same examiner and sampling reliability as did the ADOPT. However, there was no relationship between performance on that test and measured skill or practices. While the Oklahoma test is only one example of a State road test, it is typical of State road tests with respect to the behaviors it attempts to assess and the way it attempts to assess them. Like other road tests, it is highly subjective, relying upon examiners to decide what driver performances to assess, where they should be asssessed, and how to assess them. The approach is quite different from the ADOPT, which specifies the behaviors to be observed, the specific locations at which to observe them, and the criteria for deciding whether or not they are correctly performed.

There is nothing in the scoring of the Oklahoma test, or any other State test, that prevents an objective appraisal of the driver's skill and use of safe operating practices. The scoring system makes the test highly vulnerable to the influence of other factors such as the applicant's age, physical appearance, cooperativeness, or sincerity. The high reliability exhibited by the Oklahoma test could be to a great extent the result of those influences.

These factors could easily produce high sampling reliability. Any factors that related to characteristics of the applicant, rather than the applicant's performance, would be the same from one route to another. Examiners who decided that an applicant is qualified for a license on the basis of performance observed over one route are not likely to alter that decision unless something unusually significant happens on the second route. Research has shown that once an opinion about an individual is formed, it tends to influence perceptions of that individual's subsequent behavior--a phenomenon most commonly referred to as the "halo effect."

Interexaminer reliability is also susceptible to the same outside influences. Many of the perceptions that examiners have of applicants based upon characteristics of the drivers themselves are widely shared. These "stereotypes" can cause two examiners to score an applicant in the same way regardless of how the applicant actually performed.

The fact that factors other than actual performance <u>could</u> be responsible for the higher examiner and sampling reliability of the Oklahoma test does not mean that they <u>are</u> responsible for it, or that scores given to applicants are influenced by such factors. There is no way of ascertaining from the data just what was responsible for the high measured reliability of the Oklahoma test. Certainly, the applicant's actual performance can be accepted as the major determiner. However, if it is driver performance, it does not appear to be those aspects of performance that reflect the driver's skill.

Role of Skill in Safe Vehicle Operation

If the ADOPT is primarily a measure of driver skill, the next logical question is "What role does driving skill play in safe vehicle operation?"

What was probably the most comprehensive investigation into the relationship of driver factors to accident causation was that undertaken by Treat et al. (1977) at the Institute for Research in Public Safety (IRPS). In that analysis, errors associated with driver performance, the category most sensitive to driving skills, accounted for only 7% of the accidents investigated. The leading causes of accidents were lapses in use of safe driving practices, most notably those associated with visual search. These results are consonant with the opinions of highway safety specialists, who have long emphasized the criticality of safe driving practices over skills in accident prevention.

At first glance, it would appear that a measure that primarily assesses driving skill as does the ADOPT would have limited value in a licensing program if the objective of licensing is to foster safe vehicle operation. However, there are three ways most driving skills can directly contribute to the prevention of accidents.

First, while the number of accidents in which skill deficiencies were a primary cause may be small within the general population of acidents, there is evidence that they play a disproportionately higher role in accidents involving novice drivers. It is important to recognize that the purpose of a road test is primarily to assess the qualifications of new drivers. Almost all States waive a road test for drivers who already hold a valid license from another State. The test is intended primarily to see that new drivers demonstrate their ability to operate safely before being granted access to the public highways. If a license test is capable of reducing skill-related accidents, it might be expected to have a significantly larger impact upon the population of new drivers than upon drivers in general. (Unfortunately, the IRPS data are not analyzed by level of experience).

Secondly, if it is true that lapses in use of safe driving practices can result from skill deficiency among drivers taking a license test, as discussed earlier, then they can also be responsible for some of the errors that lead to accidents. Granted, among experienced drivers most of the lapses in safe driving that lead to accidents have nothing to do with skill deficiencies. However, among novice drivers, it is possible that a substantial share of the accidents that are attributed to failure to signal, failure to check mirrors, or other deficiencies in safe driving practice were at least contributed to by skill deficiencies.

Finally, one of the situations in which driving skill is likely to have its greatest impact upon safety is in the last-second avoidance of impending accidents. It is difficult to tell to what extent driving skills contribute to the avoidance of impending accidents since no records are kept on accidents that are successfully avoided. However, an analysis by Drahos and Treat (1975) disclosed that, in over half of the accidents that do occur, at least one of the drivers perceived the situation in time to avoid the accident. The fact that the drivers obviously failed to avoid it means that any attempts to evade the accident were unsuccessful. There is, of course, no way of knowing what proportion of these accidents could have been prevented if drivers had been more skillful. However, the statistics indicate that the involvement of skill in accident prevention is probably higher than the 7% of accidents directly attributed to lack of skill.

One tangential bit of evidence bearing upon the role of skill tests in accident prevention comes from an assessment of a motorcycle operation licensing program carried out by the State of California (Anderson, Ford, and Peck, 1980). This study disclosed that implementation of a new skill test resulted in a 15% reduction in accidents by applicants during the first year following administration of the test. What is of particular significance is that this is the only study in which the accident prevention value of a license test has been conclusively demonstrated. The particular test involved was administered off-street. However, it certainly establishes the importance of skill testing in accident prevention.

CONCLUSIONS

From the results of the effort described in this report, the following conclusions may be offered concerning the ADOPT:

- The ADOPT is a highly reliable measure of applicant performance. Mean scores attained by applicants from different examiners and across different routes are virtually identical. The intercorrelation of scores across examiners exceeds .8 and across routes exceeds .7. The total measurement reliability, as indicated by the correlation of scores across examiners <u>and</u> routes, exceeds .7.

- The ADOPT provides a valid measure of driving skill as independently measured through an off-street test of automobile driving skill. Both the Skills and Practices components of the ADOPT contribute to this correlation.

- A typical State road test, as represented by the Oklahoma license test, shows the same high reliability as the ADOPT. However, it evidences no relationship with an independent skill measure.

- Neither the ADOPT nor the Oklahoma road test is correlated with the use of safe operating practices by license applicants as observed unobtrusively while they are driving away from the licensing station.

REFERENCES

Anderson, J.; Ford, J. L.; and Peck, R.C. Improved Motorcyclist Licensing and Testing Project. Vol. I. Washington, D.C.: Government Printing Office, 1980. Prepared for the U.S. Department of Transportation.

Barch, A. M. "Judgments of Speed on the Open Highway." Journal of Applied Psychology, 42 (1958), 362-366.

Brown, S. L. and Bohnert, P. J. Alcohol Safety Study: "Drivers Who Die." Final Report. Baylor University College of Medicine, Houston, Texas. Contract #FH-11-6603, National Highway Safety Bureau, 1968.

Campbell, B. J. "A Comparison of the Driving Records of 1,100 Operators Selected at Random." Traffic Safety Research Review, 1958.

Cobb, P. W. Automobile Driver Tests Administered to 3,663 Persons in Connecticut, 1936-37, and the Relation of the Test Scores to Accidents Sustained. Unpublished report to the Highway Research Board, July 1939.

Crancer, A., Jr. Predicting Driving Performance with a Driving Simulator. Olympia, Wash.: Department of Motor Vehicles, 1968.

Crancer, A., Jr., and Quiring, D. L. "Driving records of persons hospitalized with suicidal gestures." Behavioral Research in Highway Safety, 1970, 1, 33-42.

Drahos, R., and Treat, J. R. An Analysis of Emergency Situations, Maneuvers, and Driver Behaviors in Accident Avoidance. Falls Church, Va.: URS/Matrix Company Systems Research Division, 1975. Prepared by the Institute for Research in Public Safety.

Dreyer, Dell R. An Evaluation of California Driver's Licensing Examination. Research Report 5. Sacramento: California Department of Motor Vehicles, 1976.

Edwards, D. S., and Hahn, C. P. "Use of Filmed Driver Behavior in the Study of Accidents: First Year Findings." American Institutes for Research, 1964.

Edwards, D. S., and Hahn, C. P., Use of Filmed Behavior in the Study of Accidents, American Institutes for Research, 1970.

Edwards, D. S.; Hahn, C. P.; and Fleishman, E. A. Evaluation of Laboratory Methods for the Study of Driver Behavior: The Relation Between Simulator and Street Performance. Washington, D.C.: American Institutes for Research, 1969.

Edwards, M. L. "An analysis of the relationship between personality factors and accident involvement." Texas Office of Traffic Safety, 1972.

Forbes, T. W.; Nolan, R. O.; Schmidt, F. L.; and Vanosdall, F. E. "Driver Performance Measurement Based on Dynamic Driver Behavior Patterns in Rural, Urban, Suburban and Freeway Traffic." Accident Analysis and Prevention, Vol. 7, pp 257-280, 1975.

Fletcher, E. D. "Visual acuity and safe driving." Journal of the American Optometric Association (1949) 20: 439-442.

Greenshields, B. D. "Driving behavior and related problems." Highway Research Record (1963) 25: 14-32.

Greenshields, B. D., and Platt, F. N. "Objective Measurements of Driver Behavior." Detroit, Mich.: SAE Automotive Engineering Congress, 1964.

Harootyan, L. D., Jr. The Effect of Experience on Longitudinal and Lateral Control Characteristics of Automobile Drivers. Master's Thesis, The Ohio University, Columbus, Ohio, 1969.

Harrington, D. M. An Evaluation of the Driving Test as an Examination Requirement for Drivers Previously Licensed in Another State. Research Report 44. Sacramento: California Department of Motor Vehicles, 1973.

Haviland, C. V., and Wiseman, H. A. B. "Criminals who drive." Proceedings of the 18th Annual Conference of the American Association for Automotive Medicine, Toronto, Canada, September 1974.

Henderson, R. L., and Burg, A. Vision and Audition in Driving. System Development Corporation, TM(L)-5297/000/00, April 1974.

Henderson, R. L., Burg, A., and Brazelton, F. A. Development of an Integrated Vision Testing Device: Phase I. Final Report No. TM(L)-4843/000/00, Systems Development Corporation, Santa Monica, CA, 1971.

Hofstetter, H. W. "Visual acuity and highway accidents." Journal of the American Optometric Association, 1976, 47, 887-893.

Kaestner, Noel. Study of Licensed Drivers in Oregon. Part II. Analysis of Traffic Involvement Records. Salem, Oregon: Department of Motor Vehicles, 1964.

Kimball, K. A.; Ellingstad, V. S.; and Hagan, R. E. An Investigation of the Acquisition of Driving Skills, Technical Report 11, Human Factors Laboratory, Department of Psychology, University of South Dakota, Vermillion, SD, 1970.

Lauer, A. R. "Motor vision." Journal of the American Optometric Association, 1937, 9, 317-323.

Lauer, A. R.; DeSilva, H. R.; and Forbes, T. W. Report to the Highway Research Board, 1939 (unpublished).

Lohman, L. S.; Legget, E. C.; Steward, J. R.; and Campbell, B. J. Identification of Unsafe Driving Actions and Related Countermeasures. Chapel Hill, North Carolina: University of North Carolina Highway Safety Research Center, 1976.

McGuire, F. L. "Personality factors in highway accidents." Human Factors (1976) 18: 433-442.

McKnight, A. James, and Adams, Bert B. Driver Education Task Analysis. Vol. 1: Task Descriptions. Washington, D.C.: Government Printing Office, 1970. Prepared for the U.S. Department of Transportation. NTIS No. PB 197 325.

McMurray, L. "Emotional stress and driving performance: The effects of divorce." Behavioral Research in Highway Safety, (1970) 1: 100-114.

McPherson, Kenard; McKnight, A. James; and Anne C. Knipper. Motorcyclist In-Traffic Test, Final Report. Prepared for the U.S. Department of Transportation, National Highway Traffic Safety Administration, Contract No. DOT-HS-7-01526, February 1978.

McRae, J. D. The Relation of Licensing Test Scores to Subsequent Driver Performance. No. 67. Chapel Hill, N.C.: University of North Carolina, 1968.

Mourant, R. R. and Rockwell, T. H. "Learning of Visual Patterns by Novice Drivers." Proceedings Federation Internationale des Societes d'Ingenieurs des Techniques de l'Automobile, Brussels, 1970.

Phillips, D. T. "Motor vehicle fatalities increase just after publicized suicide stories." Science (1977) 196: 1464-1465.

Quenault, S.W. Dissociation and Driver Behavior. Road Research Laboratory, Crowthorne, Berkshire, England, 1968.

Quenault, S.W., and Parker, P. M. Driver Behavior--Newly Qualified Drivers. Transport and Road Research Laboratory, Department of the Environment, Crowthorne, Berkshire, England, 1973.

Schmidt, C.W., Jr.; Shaffer, J. W.; Zlotowitz, H. I.; and Fisher, R. S. "Personality factors in crashes: Age and alcohol." Proceedings of the 20th Conference of the American Association for Automotive Medicine. Atlanta, GA, November 1976.

Selzer, M.; Rogers, J. E.; and Kern, S. "Fatal Accidents: The Role of Psychopathology, Social Stress and Acute Disturbance." American Journal of Psychiatry (1968) 124: 46-54.

Shinar, D.; McDonald, S. T.; and Treat, J. R. "The interaction between causally implicated driver mental and physical conditions and driver errors causing traffic accidents: An analytical approach and a pilot study." Journal of Safety Research (1978) 10: 16-23.

Silver, E. H. "Report of the Research Department of the American Optometric Association." Journal of the American Optometric Association (1936) 8: 63-69.

Tabachnick, N.; Litman, R. E.; Osman, M.; Jones, W. L.; Cohn, J.; Kasper, A.; and Moffat, J. "Comparative Psychiatric Study of Accidental and Suicidal Death." Archives of General Psychiatry, (1966) 14: 60-68.

Thorndike, R. L., The Human Factor in Accidents with Special Reference to Aircraft Accidents, Project #21-30-001, Report #1, USAF School of Aviation Medicine, Randolph Field, Texas, 1951.

Tillman, W. A., and Hobbs, G. E. "The Accident Prone Automobile Driver, A Study of Psychiatric and Social Background." American Journal of Psychiatry, CVI, 5. (1949).

Treat, J. R., et al. Tri-Level Study of the Causes of Traffic Accidents: Final Report. Washington, D.C.: Government Printing Office, 1977. Prepared by the Institute for Research in Public Safety for the U.S. Department of Transportation.

Wallace, Jean E., and Crancer, Alfred, Jr. Licensing Examinations and Their Relation to Subsequent Driving Record. Report 019. Olympia, Wash.: Department of Motor Vehicles, 1969.

Waller, J., and Goo, J. "Accident Violation Experience and Driving Test Score." Highway Research Record, 225, 1968.

Zell, John K., Driver Eye Movements as a Function of Driving Experience. Technical Report IE-16, Engineering Experiment Station, The Ohio State University, Columbus, Ohio, June 1969.

APPENDIX A

165 CANDIDATE BEHAVIORS

(with classifications assigned
by the selection panel indicated)

DRIVER/VEHICLE READINESS

- Preparing Driver
  - Wears glasses if necessary NT-V
  - Wears hearing aid if necessary NT-V
  - Is not under physical or emotional strain NT-V
  - Has not taken drugs or alcohol NT-V
- Preparing Vehicle
  - Makes sure windshield is clean NT-V
  - Tests wipers for proper operation NT-V
  - Fastens safety belts NA-L
  - Adjusts mirrors NT-V
  - Tests headlights for proper operation
    - Day NT-V
    - Night NT-V
  - Tests directional signals for proper operation
    - Day NT-V
    - Night NT-V
  - Makes sure the following are ready for safe operation
    - Tires NT-V
    - Brakes NT-V
    - Steering NT-V
  - Loads objects in car or trunk properly NT-V
  - Adjusts seat NA-L
  - Locks door NA-L
- Securing Vehicle
  - Sets brake NA-L
  - Removes key NA-L
  - Locks door NA-L
  - Positions wheels properly when parking NC

VEHICLE CONTROL

- Accelerating
  - Accelerates smoothly and evenly F
  - Shifts gears smoothly with manual shift NC
  - Selects proper gear for operation NC
- Braking
  - Releases parking brake before moving NC
  - Applies brake smoothly and evenly F
  - Releases accelerator while braking F

Steering
- Grips steering wheel correctly F
- Turns steering wheel in proper direction
  - Moving forward NC
  - Moving backward NC

VEHICLE MANEUVERING

Controlling Speed
- Achieves and maintains appropriate speed
  - Moving forward F
  - Moving backward NC
  - Going uphill NT-R
  - Going downhill NT-R
  - Entering curve F
- Downshifts when necessary to maintain speed NC
- Stops smoothly at designated point
  - Level surface F
  - Going uphill NT-R
  - Going downhill NT-R
- Maintains firm pressure on brake pedal when stopped NT-O

Controlling Direction
- Maintains straight path
  - Moving forward F
  - Moving backward NC
- Maintains curved path
  - Moving forward F
  - Moving backward NC
- Maintains path in lane change NC
- Maintains path in parallel parking NC

Gap Acceptance
- Fixed objects
  - Judges clearance
    - Between two objects F
    - From object on left F
    - From object on right F
- Moving objects
  - Entering roadway F
  - Merging NT-R
  - Passing NT-R
  - Crossing traffic F
  - Turning right into traffic F
  - Turning left into traffic F
  - Turning left across traffic F

INTERACTING WITH HIGHWAY/TRAFFIC ENVIRONMENT

- Searching
  - Ahead
    - Distance scanning
      - Looks a suitable distance ahead
        - Straight NT-O
        - Curves NT-O
      - Looks side to side F
    - Roadway
      - Signs
        - Warning signs NT-O
        - Regulatory signals NT-O
        - Informational signs NT-O
      - Signals
        - Warning signals NT-O
        - Regulatory signals NT-O
        - Informational signals NT-O
      - Pavement markings NT-O
      - Surface
        - Condition of roadway shoulder NT-O
        - Condition of roadway surface NT-O
      - Configuration
        - Curves NT-O
        - Hills NT-O
        - Ramps NT-O
    - Traffic
      - Vehicles entering roadway NT-O
      - Vehicles traveling the same direction NT-O
      - Vehicles traveling the opposite direction NT-O
      - Roadside traffic
        - Straight NT-O
        - When turning NT-O
  - Behind
    - Checks periodically NC
    - Changing direction
      - Changing lanes F
      - Merging F
      - Turning F
    - Changing speed
      - Stopping F
      - Negotiating hills NT-R
    - Backing up F
  - Rear quarter
    - Entering roadway F
    - Changing lanes F
    - Merging NT-R
  - Side to side
    - Intersections
      - With right of way
        - Blind F
        - Normal F
      - Without right of way F
    - Railroad crossing NT-R

Communicating
- Intentions
  - Changing direction
    - Turning F
    - Passing F
    - Changing lanes F
    - Entering or leaving traffic F
    - Merging F
  - Changing speed
    - Stopping NA-L
    - Turning NA-L
    - Parking NA-L
- Presence
  - Turning on headlights NT-R
  - Using emergency warning devices NT-R

Managing Speed
- Sight distance limitations
  - Roadway configuration
    - Slows for hills NT-R
    - Slows for curves F
    - Slows for blind intersections F
  - Slows during bad weather NT-R
  - Slows at night NT-R
  - Slows for sun glare NT-R
- Traction limitations
  - Slows for rain NT-R
  - Slows for snow and ice NT-R
  - Slows for gravel and debris NT-R
- Signs, signals, and markings
  - Slows at warnings NT-R
  - Does not exceed posted speed limit
    - Straight F
    - Curve/ramp F
- Traffic
  - Adjusts speed to traffic flow
    - In traffic NA-L
    - Entering traffic NA-L
  - Adjusts speed to level of congestion NA-L
  - Adjusts speed to roadside activity level NA-L

Managing Space
- Complies with roadway markings
  - Avoids stopping in crosswalk F
  - Avoids driving in restricted areas F
  - Avoids crossing or straddling painted lines F
- Selects appropriate travel lanes
  - When turning F
  - When traveling through F

- Separating from traffic
  - Ahead
    - Maintains appropriate following distance
      - When moving F
      - When stopped NC
  - Behind
    - Does not slow unnecessarily F
    - Does not stop on roadway NT-V
    - Does not back up on roadway NT-V
  - To the sides
    - Oncoming vehicles
      - Maximizes separation from oncoming vehicles F
      - Maximizes separation from oncoming left-turning vehicles F
      - Keeps wheels straight ahead while waiting to turn left NC
    - Adjacent vehicles
      - Avoids operating alongside vheicles F
      - Avoids operating in vehicles' blind spots F
    - Roadside traffic
      - Parked vehicles F
      - Pedestrians, etc. NT-R
- System Facilitation
  - Changes lanes to allow merging vehicles to enter NA-L
  - Dims lights when meeting or following vehicles NT-R
  - Obeys traffic signs and signals F

INTERACTING WITH HIGHWAY/TRAFFIC HAZARDS

- Recognizing Potential Hazards
  - Recognizes road users whose vision is obstructed NT-O
  - Recognizes road users who are distracted NT-O
  - Recognizes road users who are losing control of the vehicle NT-O
  - Recognizes road users whose path is obstructed NT-O
- Selecting Responses to Emergencies
  - Conflicts
    - Fixed objects NA-S
    - Oncoming vehicles NA-S
    - Overtaking vehicles NA-S
    - Lead vehicles NA-S
    - Intersecting vehicles NA-S
    - Converging vehicles NA-S
    - Pedestrians NA-S
  - Vehicle failures
    - Control failures
      - Steering NA-S
      - Brakes NA-S
      - Stuck Accelerator NA-S
      - Stalling NA-S
      - Tires NA-S
    - Visibility
      - Lights NA-S
      - Hood latch NA-S

Environmental problems
- Wind NT-R
- Obstacles NT-R
- Slippery spots NT-R

Executing Responses to Emergencies
- Recovers from skid NA-S
- Brakes quickly F
- Accelerates quickly NC
- Steers quickly NA-S
- Observes quickly NT-O
- Signals quickly NT-R
- Uses parking brake to stop NC

Classification Codes

F	Feasible
NC	Noncritical
NT-V	Nontestable for lack of validity
NT-R	Nontestable for lack of reliability
NT-O	Nontestable by observation
NA-L	Nonacceptable for lack of legal requirement
NA-S	Nonacceptable for reasons of safety

APPENDIX B

LIST OF EXPERT REVIEWERS

CRITICALITY TO SAFETY

Dr. Richard W. Bishop, Florida State University
Mr. Philip Cornwell, Indiana University
Mr. Rodger J. Koppa, Texas Transportation Institute
Dr. Kenard McPherson, NPSRI
Mr. Herbert Miller, NHTSA

TESTABILITY

Dr. Mark Lee Edwards, NPSRI
Dr. Margaret H. Jones, University of Southern California
Dr. A. James McKnight, NPSRI
Mr. Michael Ratz, California Department of Motor Vehicles
Mr. Michael Rudisill, Michigan State University

ACCEPTABILITY

Mr. Wayne Green, Department of Motor Vehicles, Nebraska
Mr. Newman Jackson, Retired from Texas DMV
Mr. Adam Johnson, Michigan Department of State
Maj. Howard R. Showe, Maryland Motor Vehicle Administration
Maj. Thomas Tennery, Oklahoma Department of Public Safety

APPENDIX C

SELECTION PANEL MEMBERS

Dr. Mark Lee Edwards, NPSRI
Dr. A. James McKnight, NPSRI
Dr. Kenard McPherson, NPSRI
Mr. John O'Brien, New York Department of Motor Vehicles
Mr. Michael Rudisill, Michigan Department of State
Maj. Howard Showe, Maryland Motor Vehicle Administration
Mr. Michael F. Smith, NHTSA
Maj. Thomas Tennery, Oklahoma Department of Public Safety
Dr. Stephen Versace, NHTSA

APPENDIX D

SAMPLE SCORE SHEET
(portion only)

DIRECTIONS/REFERENCE	LOCATION	PERFORMANCE CHECKS No.	CAT	YES	NO	N/A
Standard Parallel Parking Instructions	STATION LOT	1	HT	___	___	___
		2	HD	___	___	___
		3	HS	___	___	___
		4	HO	___	___	___
		5	HP	___	___	___
"At the street, turn right"	EXIT & ELM	6	GS	___	___	___
(At Elm & C) "Continue straight ahead"	ELM AT WEST C	7	BA	___	___	___
	ELM FROM WEST C TO F	8	MS	___	___	___
		9	LS	___	___	___
	ELM FROM TRAILER PARK	10	MS	___	___	___
	ENTRANCE TO CUSTOM CABINETS	11	LS	___	___	___
"At next street turn left"	(Elm at Custom Cabinets)					
	ELM AT K PLACE	12	ST	___	___	___
	("normal" speed: XX mph)	13	MT	___	___	___
		14	LT	___	___	___
"Continue straight ahead"	(After turn on K Place)					
(K Place at 2nd speed sign)	K PLACE AT JUNIPER	15	OB	___	___	___
"Next street turn left"		16	ST	___	___	___
	("normal" speed: XX mph)	17	MT	___	___	___
		18	LT	___	___	___
	JUNIPER AT CURVE	19	SC	___	___	___
		20	MC	___	___	___
(G at Ivy) "Turn right"	G AT IVY	21	GS	___	___	___
Standard RS Instructions (Speed 25 mph)	(After turn on Ivy) Alt. loc.: Next speed sign					
Signal ref.: 25 mph sign/	Dist. ref.: factory sign	22	RS	___	___	___
"Next street turn right"	(Ivy at black mailbox)					
	IVY AT D	23	OB	___	___	___
		24	ST	___	___	___
	("normal" speed: XX mph)	25	MT	___	___	___
"Next street turn left"	(D after turn)					
	D AT JUNIPER	26	BA	___	___	___
		27	OS	___	___	___
		28	RT	___	___	___