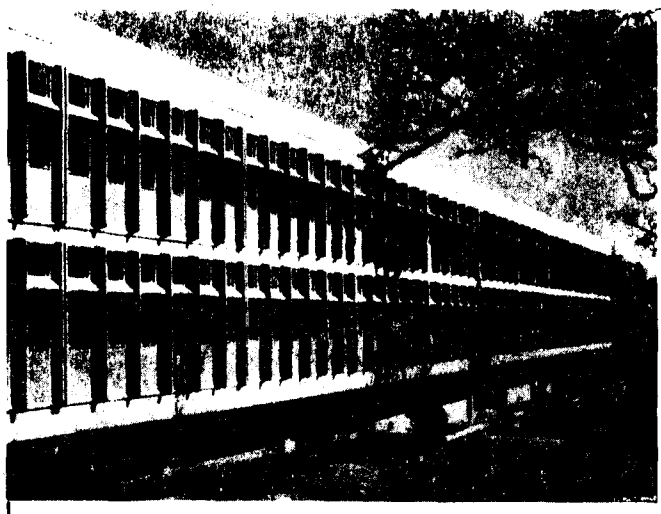
McRae, D.J. (1968). Relation of Licensing Test Scores to Subsequent Driver Performance. Chapel Hill, NC: L.L. Thurstone Psychometric Laboratory, University of North Carolina



THE RELATION OF LICENSING TEST SCORES TO SUBSEQUENT DRIVER PERFORMANCE

Douglas J. McRae

August 1968

Number 67

THE L. L. THURSTONE PSYCHOMETRIC LABORATORY
UNIVERSITY OF NORTH CAROLINA

CHAPEL HILL, N. C. 27514

ATTENTION

The enclosed report is a reprint of the original technical report which has recently gone out of print. Its content does not differ in any way from the original report. The format differs slightly due to time restrictions in the reprinting process.

We hope that this report will fulfill your interests. We appreciate your continued concern in highway safety.

THE RELATION OF LICENSING TEST SCORES TO SUBSEQUENT DRIVER PERFORMANCE[1]

A 2x3 multivariate analysis of variance was used to study the relations between test scores received by drivers at the time of licensing and future driving performance. Three groups of drivers were defined: those with no accidents or violations, those with minor violations but no accidents, and those with accidents and violations. The sex of the driver defined the second dimension of the MANOVA. The dependent variables consisted of 39 scores taken from the road test and other tests given to all driver license applicants.

No significant differences were found on the multivariate tests. However, univariate tests on the total road test score did yield significant results, and the multivariate tests for Driver Performance did yield a probability of .062. Hence, six contrasts were defined and subsets of variates that contributed most heavily to the differences defined by these contrasts were found. Discriminant function coefficients based on these reduced sets of variates were then used to construct new weighting systems for the items on the road test, and these new weighting systems were validated using an independent sample of drivers.

The conclusion of the study is that a linear composite of the items on the road test can be used to differentiate among various categories of drivers.

[1]This research was supported in part by PHS Training grant MH-8258 from The National Institute of Mental Health, Public Health Service and in part by the Highway Safety Research Center at the University of North Carolina. The author would like to thank Dr. Lyle V. Jones, Dr. Elliot M. Cramer, and Dr. B. J. Campbell for their guidance throughout the preparation of this report, and Mr. Edward A. Youngs for his work on several of the computer programs necessary for this research. A more complete description of this research is available in the author's M.A. thesis.

In general, the causes of automobile accidents may be grouped into two categories, one concerned with the environment in which a driver performs, and the other concerned with the driver himself. The first category includes all the physical characteristics of the driving situation, including the design and functioning of the automobile, the design of highways, proper placement of highway signs, etc. Much attention has been devoted in recent years to this area of highway safety. Attention also has been focused on the human variable, by the introduction of safety campaigns, through efforts to modify the driving habits of chronic accident and violation repeaters, and by developing tests with which to screen driver applicants.

Research on the human variable in highway safety may also be divided into two areas: the prediction of future driving performance, and the development or modification of driving habits (Schuster, 1966). Research on prediction of future driving performance generally deals with the search for relevant variables and combinations of variables that yield significant prediction. The research reported in this paper falls under the heading of prediction; more particularly, this study investigates the use of specific driving skills as the predictor variables.

Review of Literature. A number of different classes of variables have been investigated as possible predictors (for summaries, see Goldstein, 1961 and 1964). Physical characteristics of potential drivers, personality and attitude characteristics, previous driving record, situational variables (effects of fatigue, effects of alcohol), and various types of personal data (age, sex, etc.) have all been shown to yield some differentiation between different classes of drivers, but predictive ability has been limited. Other variables, such as reaction time measures, psychomotor

task variables, sensory perceptual tests, and cognitive measures have yielded less promising results.

Little research has been done on the class of predictors investigated in this study, actual driving skills (AMA, 1966, p.265). Three World War II Army studies are summarized by Goldstein as yielding non-significant validity coefficients of .01, .03, and -.04, prompting the comment "This lack of relationship between road tests and accidents makes road tests extremely dubious as measures of safe driving." In another Army research project, Uhlaner (1966) concludes that "... present day selection procedures in public licensing of drivers can make only dubious contribution to the accident reduction problem."

Campbell (1958) compared the licensing test scores of 1100 motor vehicle operators involved in fatal accidents with 1100 operators selected at random. Only 3 items (parallel parking, second left turn signal, and third left turn signal) of 36 significantly differentiated between the two categories of drivers. In view of such results, A.R. Lauer of the Driving Research Laboratory at Iowa State University has suggested that the skills portion of the driver licensing examination be limited to a demonstration of proficiency at parallel parking (AMA, 1966, p.265).

Methods

This study attempts to relate two sets of variables: driver's licensing test scores gathered during the licensing procedure and subsequent driving performance as measured by accidents and violations. Information on both sets of variables is kept by the Motor Vehicles Department (Raleigh, North Carolina) for all current North Carolina drivers who were originally licensed in North Carolina.

The study can be divided into two parts. First, there is an attempt to discover whether there are any significant

differences between drivers with distinguishing kinds of driver performance records (good and bad drivers) on the test scores assigned during the licensing procedure. Second, there is an attempt to use these differences to predict the performance records for an independent sample of drivers.

The analysis done for the first part is a multivariate analysis of variance (MANOVA) in which the variates are the test scores from the licensing procedure. The major factor (or dimension) of interest is that of driver performance. Variables controlled for in this section of the study include age, sex, amount of driving, time of licensing, and time since licensing.

The second part of the study is described in the Validity section. The discriminant function coefficients associated with the significant roots on the driver performance dimension were used to construct new weighting systems for the road test items. An independent sample (validity sample) of drivers was then scored on these new scoring systems in an attempt to validate the developed systems. The developed scoring systems were also compared to the existing scoring system.

Controlled Factors. Five factors or variables were controlled in this study either by inclusion in the dimensions of the analysis of variance or by limitation of the sample selected. The five factors were:

1. Time of Licensing. Although there have been no major changes in the method of licensing drivers in North Carolina for the past 20 years, minor changes have occured and accumulated over time. For example, the parallel parking item on the road test was eliminated about five years ago. For this reason, it was decided to limit the study to those drivers who received their licenses during a relatively short

specified period of time. This period of time was June 26, 1964 to May 31, 1965.

2. Time Since Licensing. The further in time one gets from the licensing procedure, the less predictive value the test scores are likely to have. Hence, only the first two years after licensing were considered for each driver in the study.

3. Amount of Driving. This variable is of great import: a driver with one or two minor violation citations, if he drives often, may be a better driver than one who drives very little and has a clear record. Unfortunately, information on exposure to driving was not available. Some attempt was made to control this variable in the definition of the driver performance categories (see below); no claim is made that this variable has been adequately controlled in this study.

4. Sex. Previous studies (Campbell, 1958; Schuster, 1966; Levonian, 1967) have found that sex is related to driver performance. This variable was included as one dimension in the MANOVA design.

5. Age. Previous studies also have indicated that age is related to driver performance. As a control for age, the study was limited to individuals licensed between the ages of 16 and 20.

Definition of the Driver Performance Categories. A natural way of defining good and bad drivers might be to classify drivers with clear records as "good" and drivers with a certain number of accidents and/or violations as "bad". As pointed out above, this categorization suffers from a lack of control on the amount of driving, particularly in the "good" category where a clear record may indicate a good dri-

ver, or alternatively, a driver who drives very little.

An approach toward control of this variable was made by assuming that persons with minor violations but no accidents represent better drivers than do persons involved in accidents in which they have also been charged with a violation. Hence, two categories for driver performance were defined with some small confidence that the amount of driving variable was taken into account: first, an "acceptable" category consisting of drivers with a small number (not more than two) of minor violations but no accidents, suspensions, or revocations; and second, an "unacceptable" category consisting of drivers involved in a certain number of accidents in which they have also been charged with a violation (hereafter called "at-fault" accidents). A third category of drivers, those with clear records, was also included in the design.

The "acceptable" category of drivers was defined to consist of drivers with minor violations but no accidents. The restriction "minor" was operationalized by defining any violation assigned a point value[2] of four or greater a "major" violation and any violation assigned a point value of three or less a "minor" violation.

The "unacceptable" category of drivers was originally defined to include drivers involved in two or more at-fault accidents. It was anticipated that there would not be an adequate number of drivers meeting this criterion; hence, an alternate criterion was developed, to include drivers involved in one at-fault accident and, in addition to the violation associated with the accident, at least one other major violation. Approximately 45 percent of the accident category drivers used in the study actually met the requirements of

[2] A prespecified number of points is assigned to a North Carolina driver's record for each violation citation he receives. Accumulation of points over a period of time may lead to suspension or revocation of the driver's license.

the alternate definition.

Summary of Design. In summary, the design for this study is a 2x3 (sex x driver category) multivariate analysis of variance. Only those persons licensed between June 26, 1964 and May 31, 1965 between ages 16 and 20 were elgible for the study. Admission to the driver performance categories was as follows:

A: Clear records in the two years following original licensing;

B: At least one but not more than two minor violations, together with no accidents, suspensions, or revocations, in the two years following original licensing;

C: Involvement in at least two "at-fault" accidents in the two years following original licensing; or involvement in one "at-fault" accident and, in addition to the violation associated with the accident, at least one other major violation in the two years following original licensing.

Choice of Variables. The variates for the study were chosen from the 47 scores on the items of the road test and the three total scores -- scores on the knowledge of signs tests, the knowledge of rules test, and the total road test score. Since the total road test score is a linear combination of the 47 items of the road test, it cannot be included in the multivariate analyses; separate univariate analyses were performed with this score. On the road test, each turn is scored separately for lane, turn, speed, and signal. By combining the first three of these under the general heading "turn", the number of variates is reduced to 37 from the road test plus two total scores (signs test and rules test). A list of the variates used in the study is given in Table 1.

Data Collection. Two programs were written to select the subjects for this study from records on magnetic tape

Table 1: Variates

		Points Deducted		
		Good	Fair	Bad
1.	Quick stop	0	10	20
2.	Backing 50 feet	0	4	8
3.	Hand brake stop	0	3	6
4.	Turn about	0	4	8
5.	Stop and start on grade	0	4	8
6.	Shifting going down	0	3	6
7.	Posture	0	2	4
8.	Clutch	0	2	4
9.	Attention	0	3	6
10.	Distraction	0	3	6
11.	Keeping in lane	0	2	4
12.	Following	0	2	4
13.	Overtaking	0	1	2
14.	Being overtaken	0	1	2
15.	Right of way	0	1	2
16.	Use of horn	0	1	2
17.	Time (compared to normal)	0	6	12
18.	First start	0	1	2
19.	Second start	0	1	2
20.	First approach to corner	0	1	2
21.	Second approach to corner	0	1	2
22.	First slow sign	0	1	2
23.	Second slow sign	0	1	2
24.	First stop sign	0	2	4
25.	Second stop sign	0	1	2
26.	First traffic signal	0	2	4
27.	Second traffic signal	0	1	2
28.	First left turn	0	3	6
29.	First left turn: signal	0	2	4
30.	Second left turn	0	3	6
31.	Second left turn: signal	0	2	4
32.	Third left turn	0	3	6
33.	Third left turn: signal	0	2	4
34.	First right turn	0	3	6
35.	First right turn: signal	0	1	2
36.	Second right turn	0	3	6
37.	Second right turn: signal	0	1	2
38.	Signs test	Range:	0 to 30	
39.	Rules test	Range:	0 to 30	
40.	Total road test	Range:	0 to 30	

at the Motor Vehicles Department. The first program examined the accident and violation record (and appropriate demographic information) for each driver licensed between June 26, 1964 and May 31, 1965. For each driver who met all criteria for admission into one of the cells in the design, the name, license number, and cell identification code were written onto another magnetic tape. The second program randomly selected approximately 600 drivers per cell from the pool of drivers determined by the first program. The sample was randomly sorted down to 450 drivers per cell and the data collection proceeded to its second stage, the recording of test scores from the licensing procedure onto punched cards.

For each of the 450 drivers per cell, the original applications were pulled from the master files at the Motor Vehicles Department. A preliminary check of this information indicated that a small percentage of the drivers did not have complete scores on all sections of the licensing procedure. Elimination of thesedrivers left 420 to 435 per cell. At random 300 of these were chosen for the primary MANOVA analyses. The rest were saved to form the pool of drivers from which the validity sample was selected.

Sample Size. The original design for this study called for an n of 300 per cell (N=1800). However, only 29 females were found to meet the criteria for admission into the Accident category. In addition, after the data were collected, a percentage of the data were found to be inconsistent, in that the total road test score was not a simple sum of the variates on the road test. Elimination of these drivers left the following non-orthogonal MANOVA design:

	Clear Record	Minor Violation	Accident
Male	263	256	257
Female	266	253	24

All analyses reported in the Results section of this report are based on this sample.

The original design for the validity sample called for an n of 100 per cell (N=600). However, no drivers remained for the female Accident category, and a percentage of the selected drivers for the other cells has the inconsistency described above. Elimination of these drivers left the following design for the validity study:

	Clear Record	Minor Violation	Accident
Male	85	88	82
Female	94	89	0

All analyses reported in the Validity Study section of this paper are based on this sample.

Analyses. The major analysis was the multivariate analysis of variance on the 39 variates listed in Table 1. The major hypotheses of this study involve differences among various levels on the driver performance factor. Since there are two degrees of freedom, tests may be obtained for a number of different contrasts. In all, three sets of two independent contrasts were obtained:

	Category:	Clear Record	Minor Violation	Accident
Set I:	Contrast 1:	0	1	-1
	Contrast 2:	2	-1	-1
Set II:	Contrast 1:	1	0	-1
	Contrast 2:	-1	2	-1
Set III:	Contrast 1:	1	-1	0
	Contrast 2:	-1	-1	2

These contrasts may be interpreted as direct tests for differences between the categories specified: i.e., Contrast 1 of Set I is a test for differences between the Minor Violation group and the Accident group on the 39 variates, Contrast 2 of Set I tests for differences between the Clear Record group and the average of the Minor Violation and Accident groups, and so on. These sets of contrasts are not independent, and hence, caution must be taken in their interpretation. It was felt that the inclusion of a validity sample in the design of this study would serve as a precautionary measure against overinterpretation of these individual degree of freedom tests.

Several contrasts seemed to have more subjective importance than others. Particularly, Contrast 1 of Set I singles out the differences between the Motor Violation category and the Accident category, and hence makes maximum use of the control for amount of driving. Contrast 1 of Set II looks at the differences between the Clear Record category and the Accident category and hence is a direct test of "no record" vs. "record" drivers. Finally Contrast 2 of Set III tests the differences between the average of the Clear Record and Minor Violation categories and the Accident category, and hence tests all drivers of the "good" or "acceptable" class against the "unacceptable" drivers.

All of the above analyses were performed using all 39 variates listed in Table 1. Since not all 39 variates contribute equally to the differences between categories of drivers, an attempt was made to identify subsets of variates that contributed most heavily to these differences. Subsets of variates were found for the overall Category analysis, and for each individual degree of freedom analysis. This was done using analyses of covariance: for each analysis, a subset of variables was found such that, when used as co-

variates, the rest of the variates add essentially nothing to the differences between groups, and hence may be discarded (Rao, 1965). Variates yielding univariate probability levels of 0.15 or less, or unusually large discriminant function coefficients (see immediately below), were included in these subsets as variates to be kept for further analysis.

As part of the multivariate analysis of variance, the discriminant function coefficients associated with all roots significant at the .15 level were obtained. These coefficients, properly weighted, can be applied to the raw scores from the road test to form weighted composite scores that optimally discriminate among the various categories of drivers. These weighted composite scores are of central interest in the validity study.

Finally, several variables were recorded on the data cards for each driver but were not included in the primary analyses. Included under this heading is information on race, on the model year of the car in which the road test was taken, on the condition of the car, and on the form of the rules test administered. Appropriate analyses of variance were performed on these variables.

Results

The total number of drivers receiving their license in North Carolina between June 26, 1964 and May 31, 1965 was 134,327. Of this number, 70,889 were under age 21 at time of licensing. Drivers who failed to meet criteria specified for the three driver categories were discarded. Population totals for the cells in the design appear in Table 2.

Results of the Primary Analyses. The analyses of variance for this study used the MANOVA program described by Clyde, Cramer, and Sherin (1966). This program yields an exact least squares solution for the non-orthogonal case.

Table 2: Population estimates and percentages including only those under 21 at the time of licensing

	Clear Record	Minor Violation	Accident	Discarded	
Males	13,412	4,871	588	21,084	39,955
	(18.9)	(6.9)	(0.8)	(29.7)	(56.4)
Females	22,340	1,377	28	7,189	30,934
	(31.5)	(1.9)	(0.0)	(10.1)	(43.6)
	35,752	6,248	616	28,273	70,889
	(50.4)	(8.9)	(0.8)	(39.9)	(100.0)

The overall multivariate tests for Sex X Category, for Category, and for Sex are presented in Table 3. The "Root 2" line indicates the residual test; i.e., the test based on the differences remaining after the largest root has been extracted. The multivariate tests are exact. Note that none of the multivariate results are significant at the .05 level. Nevertheless, since this study was exploratory in nature, and since the multivariate results for the Category analysis were significant at the .10 level, it was decided to continue the study according to the design presented in the Methods section despite the non-significance of the primary results. Hence, the multivariate tests for each of the six individual degree of freedom contrasts were found; these results are presented in Table 4.

In addition to the multivariate tests, corresponding univariate tests were performed on the total road test score. Recall that this score is a linear composite of the items on the road test. The results of these univariate analyses (overall and individual contrasts) are presented in Table 5. The fact that the overall test and a number of the contrasts are significant indicates promise for differentiation among categories of drivers using a linear composite of the items on the road test.

Finally, the category means, adjusted for the Sex effect, for the 39 variates in the multivariate analysis and for the total road test score are presented in Table 6. These means will be needed for some of the results discussed in the Validity section.

Reduction of the Number of Variates. As explained in the methods section, not all variates contribute equally to the multivariate results presented in Tables 3 and 4. Rather, it is the general case that a small subset of the variates account for a large part of any differences found. Hence,

Table 3: Primary Multivariate Results

	df	P
Category X Sex		
Roots 1 and 2	78/2550	.132
Root 2	39/1275	.315
Category		
Roots 1 and 2	78/2550	.062
Root 2	39/1275	.244
Sex		
Root 1	39/1275	.093

Table 4: Multivariate Results on Individual Degree of Freedom Tests

	df	P
Minor Violation vs. Accident	39/1275	.251
Clear Record vs. average of Minor Violation and Accident	39/1275	.085
Clear Record vs. Accident	39/1275	.193
Minor Violation vs. average of Clear Record and Accident	39/1275	.130
Minor Violation vs. Clear Record	39/1275	.054
Accident vs. average of Clear Record and Minor Violation	39/1275	.274

Table 5: Univariate results for the total road test score

	df	p
Category X Sex	2/1313	.021
Category	2/1313	.019
Sex	1/1313	.001

Contrasts:

	df	C X S	Category
Minor Violation vs. Accident	1/1313	.006	.008
Clear Record vs. average of Minor Violation and Accident	1/1313	.067	.111
Clear Record vs. Accident	1/1313	.011	.012
Minor Violation vs. average of Clear Record and Accident	1/1313	.017	.066
Clear Record vs. Minor Violation	1/1313	.662	.874
Accident vs. average of Clear Record and Minor Violation	1/1313	.006	.005

Table 6: Category means adjusted for Sex

	Clear Record	Minor Violation	Accident
Quick stop	.756	.609	.733
Backing	2.079	2.107	2.289
Brake stop	.051	.094	.113
Turn about	2.509	2.304	2.212
Stop and start	.586	.772	.780
Shifting	.057	.024	.039
Posture	.257	.232	.322
Clutch	.201	.255	.342
Attention	.578	.741	.718
Distraction	.094	.106	.090
Keeping in lane	.654	.715	.775
Following	.045	.027	.067
Overtaking	.006	.014	.010
Being overtaken	.017	.026	.012
Right of way	.083	.138	.150
Use of horn	.062	.073	.053
Time	.816	.672	.981
First start	.149	.130	.170
Second start	.066	.073	.078
First approach	.308	.303	.341
Second approach	.178	.153	.199
First slow sign	.125	.206	.146
Second slow sign	.093	.084	.106
First stop sign	.246	.265	.283
Second stop sign	.072	.096	.081
First traffic signal	.094	.118	.144
Second traffic signal	.047	.045	.029
First left turn	.729	.725	.772
First left: signal	.488	.523	.451
Second left turn	.737	.572	.797
Second left: signal	.367	.401	.323
Third left turn	.635	.562	.612
Third left: signal	.352	.314	.313
First right turn	.648	.710	.715
First right: signal	.195	.198	.174
Second right turn	.639	.542	.679
Second right: signal	.199	.234	.222
Signs test	3.838	3.705	4.692
Rules test	18.276	18.503	19.958
Total road test	15.218	15.162	16.323

for the overall Category result, and for each individual degree of freedom contrast, an attempt was made to reduce the number of variates to a relatively small subset that accounts for a relatively large portion of the differences between the groups of drivers in question. This was done using the analysis of covariance procedure described earlier. The multivariate and univariate results, and standardized discriminant function coefficients for the selected subsets of variates for the overall Category test and for each individual degree of freedom tests are presented in Table 7. The discriminant function coefficients are normalized, and hence must be multiplied by the reciprocal of the raw score standard deviation before they can be applied to raw score data.

Additional Analyses.

Rules test. The rules test was significant on the driver performance category dimension for the overall Category analysis for several of the individual degree of freedom analyses (see Table 7). There are six different forms, supposedly parallel, for the rules test: five written forms and one oral form. By including the form of the test as the third dimension in a three-way analysis of variance, a further analysis of the rules test can be made. The results of this analysis are presented in Table 8. The main effect for Form is significant. Post hoc contrasts indicated that Form 3 was significantly easier than the other forms.

Race. Information on the Race of the examinees was recorded on the data cards. To see whether there were any differences between the Races on the total road test scores, Race was added as a dimension in a three-way analysis of variance. Results are presented in Table 9. Due to the non-significance of the results, no further analyses were made on this variable.

Table 7: Multivariate analyses on reduced sets of variates

	p	SDFC	
Overall Category: Roots 1 and 2:		**p = 001**	
Root 2:		**p = 006**	
Stop and start	145	400	033
Clutch	051	364	321
Attention	131	353	-024
Right of way	029	475	150
First slow sign	017	477	-314
Second left turn	007	-319	582
Signs test	109	029	409
Rules test	022	257	434
Minor Violation vs. Accident: Multivariate p = 001			
Posture	111	281	
Following	106	308	
Time	056	344	
First slow sign	100	-320	
Second left turn	006	496	
Second right turn	092	271	
Signs test	041	302	
Rules test	021	365	
Clear Record vs. average of Minor Violation and Accident: Multivariate p = 001			
Brake stop	092	238	
Turn about	061	-284	
Stop and start	055	358	
Clutch	025	390	
Attention	058	288	
Keeping in lane	134	313	
Right of way	008	435	
First slow sign	061	339	
Rules test	043	351	

Table 7: continued:

	p	SDFC
Clear Record vs. Accident:	**Multivariate**	**p = 001**
Brake stop	136	276
Turn about	096	-270
Stop and start	140	277
Clutch	015	425
Keeping in lane	136	370
Right of way	028	426
Signs test	076	309
Rules test	007	446
Minor Violation vs. average of Clear Record and Accident:	**Multivariate**	**p = 001**
Following	119	322
Time	063	342
First slow sign	010	-518
Second left turn	002	584
Second right turn	056	304
Signs test	124	274
Clear Record vs. Minor Violation	**Multivariate**	**p = 001**
Turn about	150	277
Stop and start	076	-363
Attention	054	-324
Right of way	025	-409
First slow sign	005	-541
Second left turn	012	488
Accident vs. average of Clear Record and Minor Violation:	**Multivariate**	**p = 001**
Posture	135	309
Clutch	033	435
Time	109	332
Second left turn	059	395
Signs test	038	354
Rules test	007	483

Decimal points omitted.
SDFC: Standardized Discriminant Function Coefficients

Table 8: Form of rules test analysis

	df	p
Form	5/1278	.011
Sex X Form	5/1278	.413
Category X Form	10/1278	.176
Sex X Category X Form	10/1278	.643

Table 9: Race analysis

Race	1/1297	.237
Sex X Race	1/1297	.691
Category X Race	2/1297	.618
Sex X Category X Race	2/1297	.978

Table 10: Condition of car analysis

Condition of car	2/1299	.271
Sex X Condition	2/1299	.473
Category X Condition	4/1299	.281
Sex X Category X Condition	4/1299	.413

Table 11: Model year of car analysis

Sex	1/1313	.001
Category	2/1313	.718
Sex X Category	2/1313	.064

Condition of Car. The examiner records the condition of the car in which the road test is taken. By including this variable as one of the dimensions of a three-way analysis of variance, it may be tested on the total road test score. Results are presented in Table 10. Once again, the results were not significant, and no follow-up analyses were done.

Model Year of the Car. In addition to the condition of the car, the model year is also recorded. To analyze this variable, it was used as the dependent variable in a two-way analysis of variance. The results of this analysis are presented in Table 11. The Sex dimension is significant; the means for this dimension adjusted for Category (1960.182 for females and 1959.404 for males) indicate that females take the road test in significantly newer cars than do males.

Validity Study

The results reported above were based on a sample of size 1319. During the data collection, the appropriate information was recorded for an additional 438 drivers. These data were used for the Validity Study.

Development of the Composite Scores. As part of any multivariate analysis of variance, the discriminant function coefficients associated with each root may be obtained. These coefficients, when applied to the actual data, yield a composite score that maximally discriminates between groups for the data in question. For the validity study, the discriminant function coefficients obtained in the main analysis were applied to the raw data of the validity sample to develop composite scores for the validity sample.

The MANOVA program used for this study prints standardized discriminant function coefficients associated with all roots significant at the .15 level. The coefficients used to develop the composite scores for the validity study were pre-

sented in Table 7. Note that these coefficients are based on the analyses of reduced sets of variates; hence, the composite scores are based on only a small percentage of the variates, those found to contribute heavily to the differences between categories.

Composite scores were obtained for each contrast and for both roots of the overall Category analysis. These eight scores, together with the total road test score for each validity sample driver, provided the data for the validity analysis.

Analysis and Results. The total road test score was analyzed in a 2x3 univariate analysis of variance with one missing cell. The results of this analysis are presented in Table 12. Note that the probability level for Category ($p<.174$) does not reach the same level of significance obtained in the main analysis (where $p<.019$).

Two composite scores were computed based on the discriminant function coefficients from the overall Category analysis. The results of the multivariate analysis of the two composite scores using the validity sample are given in Table 13. The fact the the second composite score yields better differentiation among categories of drivers in the validity sample than does the first composite score can only be due to randon fluctuations between the main sample and the validity sample. Neither composite score is significant at the .05 level; hence, it is concluded that the discriminant weights identified in the overall Category test of the primary analyses are not confirmed by the validity study.

Composite scores were also computed for each driver in the validity sample for each of the six contrasts. These six scores were each submitted to a univariate analysis of variance. Since each of these scores was based on the discriminant function coefficients of an individual degree of freedom contrast, the same contrasts were used in the validity analysis.

Table 12: Road test analysis for validity sample

	df	p
Sex	1/433	.208
Category	2/433	.174
Sex X Category	1/433	.948

Table 13: **Multivariate results on the two composite scores derived from the overall Category analysis**

Multivariate test:	F	p	
Roots 1 and 2 :	1.947	.101	
Root 2 :	0.736	.391	
Univariate tests:	**F**	**p**	**SDFC**
Overall score 1	1.130	.324	.553
Overall score 2	2.579	.077	.874

SDFC: Standardized Discriminant Function Coefficients

Table 14: **Contrast results for the composite scores**

	df	F	p
Minor Violation vs. Accident	1/433	0.753	.386
Clear Record vs. average of Minor Violation and Accident	1/433	5.750	.017
Clear Record vs. Accident	1/433	8.359	.004
Minor Violation vs. average of Clear Record and Accident	1/433	0.001	.975
Clear Record vs. Minor Violation	1/433	0.198	.657
Accident vs. average of Clear Record and Minor Violation	1/433	3.912	.049

The results of these analyses are presented in Table 14. Evidence for the validity of the primary analysis is given by the probability levels obtained for three of the composite scores: Clear Record vs. the average of Minor Violation and Accident, Clear Record vs. Accident, Accident vs. the average of Clear Record and Minor Violation. Note that these three contrasts form an intuitively meaningful set, centering around the differences between the Clear Record category and the Accident category.

Graphic Representation. The results of the previous section indicate that different categories of drivers can be differentiated using a weighted composite score of the items on the road test. Despite the statistical significance of these results, the differences between the categories are not great.

To gain insight concerning the differentiation that this study has identified, the distributions of scores for the validity sample Clear Record category and Accident category were plotted for the Clear Record vs. Accident composite score (Figure 1) and for the total road test score (Figure 2). The differentiation between the two categories is not readily apparent for either score. Nevertheless, if one uses a cutting point suggested by the adjusted means of the main analysis, the Clear Record vs. Accident composite score yields the following 2x2 Chi Square result:

	Below	Above
CR	.64	.36
Ac	.49	.51

$x^2 = 4.577 \quad (p<.05)$

The total road test score does not yield a significant Chi Square result:

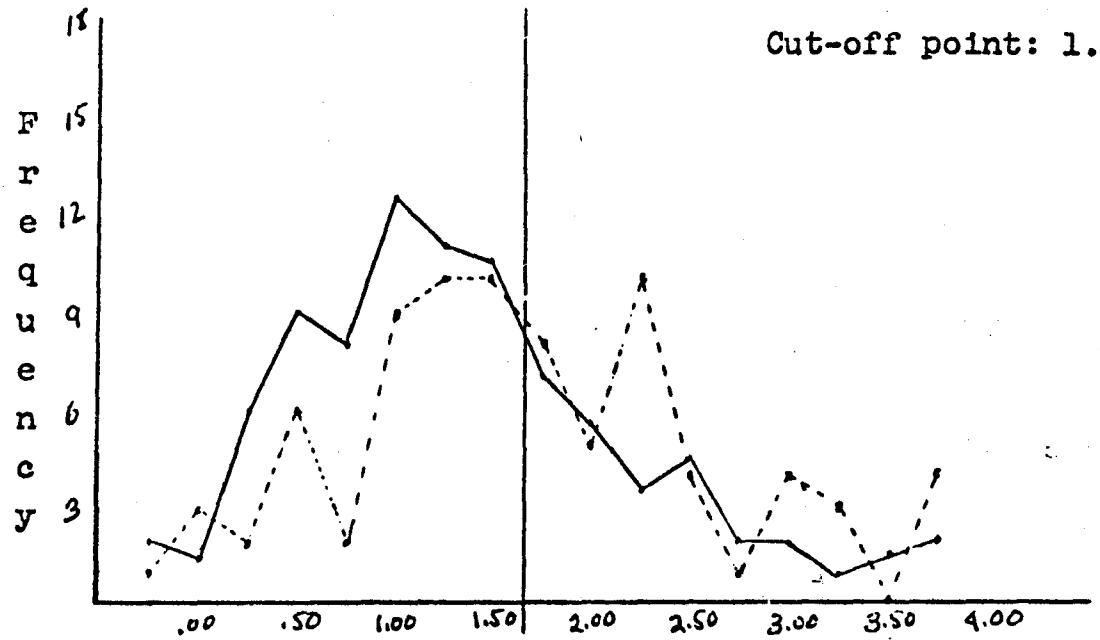


Figure 1: Clear Record (solid) and Accident (dotted) categories plotted for the CR vs. Ac composite score.

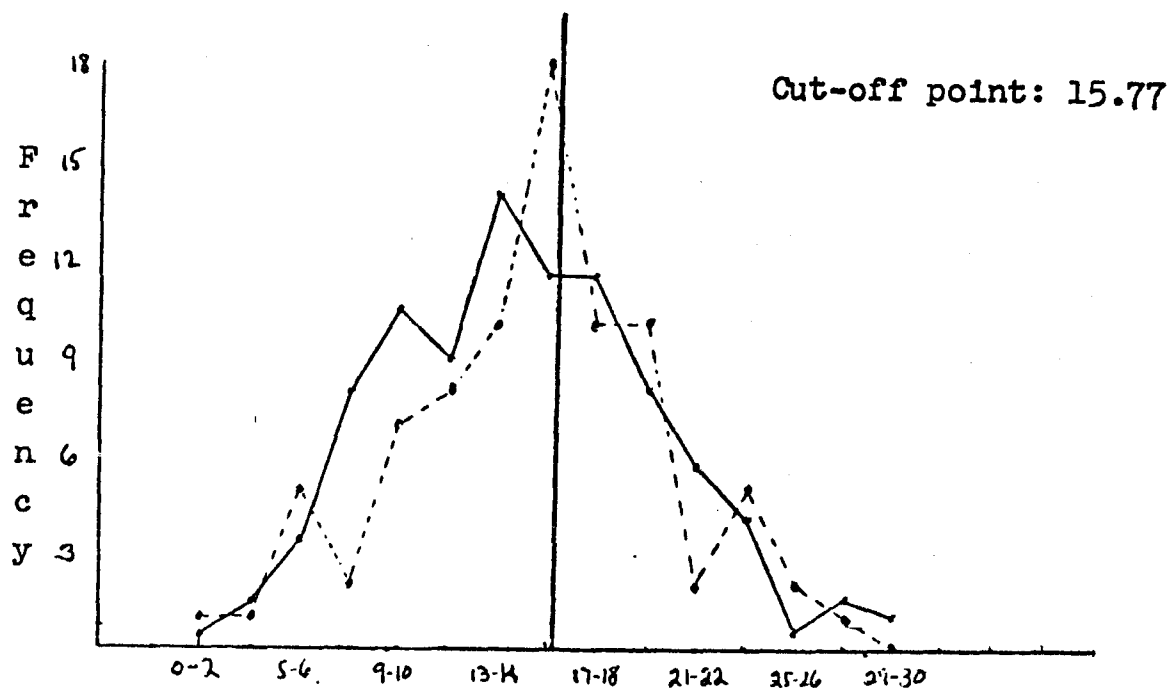


Figure 2: Clear Record (solid) and Accident (dotted) categories plotted on the total road test score.

	Below	Above
CR	.60	.40
Ac	.55	.45

$\chi^2 = 0.501$ (NS) .

These analyses confirm the finding that the composite score does differentiate between the Clear Record category and the Accident category in the validity sample, whereas the total road test score currently in use fails to differentiate.

New Weighting System. One of the major purposes of this research was to develop a new set of weights to apply to the individual maneuvers of the road test. The suggested new set of weights are given in Table 15. These suggested weights were chosen to correspond to the results presented in Tables 7 and 14; those maneuvers heavily weighted on the three contrasts receiving supporting evidence from the validity study were weighted 10 or 8, those not appearing in the reduced sets were weighted 2, etc. A number of weights were adjusted on other bases; i.e., second and third turns were more heavily weighted than first turns, and so on.

Discussion and Conclusion

The results of this study indicate that drivers with varying types of performance records do exhibit differing patterns of driving skills as measured by the licensing examination. Whether these differences can be used successfully to predict future driver performance records is an issue unresolved.

Limitations in the Design of the Study. Several shortcomings in the design of the study should be noted:

1) Exposure. As stated in the Methods section, one variable important for a study of this type is the amount of driving that a person does. The criterion for admission

Table 15: Suggested weighting system

	Good	Fair	Bad
Quick stop	0	3	6
Backing 50 feet	0	2	4
Hand brake stop	0	4	8
Turn about	0	4	8
Stop and start	0	4	8
Shifting going down	0	1	2
Posture	0	3	6
Clutch	0	5	10
Attention	0	3	6
Distraction	0	1	2
Keeping in lane	0	5	10
Following	0	2	4
Overtaken	0	1	2
Being overtaken	0	1	2
Right of way	0	5	10
Use of horn	0	1	2
Time (compared to normal)	0	2	4
First start	0	2	4
Second start	0	1	2
First approach	0	1	2
Second approach	0	1	2
First slow sign	0	4	8
Second slow sign	0	3	6
First stop sign	0	1	2
Second stop sign	0	1	2
First traffic signal	0	1	2
Second traffic signal	0	1	2
First left turn	0	3	6
First left: signal	0	1	2
Second left turn	0	4	8
Second left: signal	0	1	2
Third left turn	0	4	8
Third left: signal	0	1	2
First right turn	0	3	6
First right: signal	0	1	2
Second right turn	0	4	8
Second right: signal	0	1	2
Signs test	0	4	8
Rules test	0	5	10

Exclusing Signs test and Rules test, there are 174 possible points to be deducted. The present scoring system has 174 possible points to be deducted.

into the Accident category was a certain number of accidents and associated violations. A person who drives quite often exposes himself to accident and/or violation opportunities more than a person who seldom drives. Likewise, a person may have a clear record simply because he seldom drives, and despite the fact that evaluation of his driving skills indicate that he is not a good driver.

Some effort was made to control this variable in the present study by definition of two categories of acceptable drivers. The first category was drivers with clear records, and, it was assumed, drivers who had clear records either because they were good drivers or because they did not drive often. The second category was drivers with minor violations, drivers who had relatively acceptable records despite the assumption that, as a group, they had greater driving exposure than the first category. Hence, it is hypothesized that the major difference between these two categories was the driving exposure variable.

The results of the study did not show clear differences between these two groups. This result neither verifies nor contradicts the assumption that the minor violation group had greater driving exposure than the clear record group; rather, it says that differences between the groups on the road test are not statistically significant.

The driving exposure variable also has implications in conjunction with the Sex factor in this study. The drivers considered in the study were under 21, and it may be hypothesized that for this age group males have greated driving exposure than do females. It may also be hypothesized that females, when they do drive, stay out of trouble more than do males. Both of these hypotheses are consistent with the results presented in Table 2, where the proportion of males who fall into the Minor Violation and Accident categories is far greater

than would be expected by chance.

Due to the importance of this variable, it is recommended that future studies attempt to control it in a more systematic fashion than does the present study. Two indices that might be considered are miles driven per given period of time, and number of times driving per given period of time. Choice of an index should depend on the characteristics of the group of drivers studied. Finally, it should be noted that investigations based on professional drivers (truck drivers, bus drivers, etc.) do control for the exposure variable, but generalization from this type of driver to drivers in general may not be valid.

2) Scale of Measurement. A second methodological limitation to this study was the scale of measurement used for each item on the road test. Recall that each item was scored on a three-point scale (Good, Fair, Bad) with zero points deducted for a Good rating, and a variable number of points deducted for a Fair or Bad rating on any given item. The vast majority of the examinees received good ratings for most of the items; hence, the item distributions were heavily skewed, and the assumption of multivariate normality was not met.

Not much is known concerning the violation of assumptions for multivariate analysis of variance (Jones, 1966). Transformations for the data were considered, but the data were so skewed (more than 80 percent scored Good for most items) that the transformed data would probably have been just as badly skewed. Other possible solutions lie either in collecting data on a scale of measurement that is likely to meet the multivariate normality assumption, or in analyzing the data by a non-metric multivariate technique that does not require the multivariate normality assumption (Bock, 1966).

3) Restricted Population. A third limitation was the fact that the study does not deal with the population of all possible drivers. Theoretically speaking, it should, since

we would like to be able to predict the future driving performance of any possible driver. The population from which the sample was drawn for this study was limited to those who had already passed the licensing examination. Although this restriction should be noted, it cannot be rectified within the context of the present study, because the data required to categorize the applicant who fails do not exist.

The above discussion does not imply that the data on the applicant who fails cannot be used in the context of another study. Unfortunately, these data are not kept under the present record-keeping system in North Carolina, but data could be collected to compare the patterns of test scores of the failure group against the patterns of test scores for the Clear Record group, the Minor Violation group, and the Accident group. In addition, the number of times a license applicant takes the road test may be a valuable piece of information; it may be, for example, that differing sets of weights should be assigned to the road test items depending on whether the applicant is taking the test for the first, second, etc., time. If data on drivers who fail the test oneor more times can be made available, research could be designed to answer these types of questions.

Discussion of the Results. The differences between the categories of drivers identified by this study are small but statistically significant. Furthermore, the differences between the Clear Record category and the Accident category are consistent enough to be validated by an independent sample of drivers.

As shown in Figure 1 above, the distributions of the categories of drivers overlap considerably. Nevertheless, using the cutoff points suggested by the adjusted means from the main analysis, statistical significance is again obtained for the composite scoring system developed in this study.

The suggestion might be made that the cutting point suggested in Figure 1 would be optimal for discrimination and therefore should be used as a basis for pass or fail on the licensing examination. For argument sake, let's say that using this cutoff, the true situation in the population of all possible drivers is the following:

	Pass	Fail
Future good	65 percent	35 percent
Future bad	50 percent	50 percent

The above situation is close to the results given in Figure 1. Using the suggested cutoff would then require failure of 35 percent of all future good drivers in order to fail 50 percent of future bad drivers. Recall that Table 2 shows only 1.69 percent (616 of 36,368) drivers are identified as accident category drivers out of all drivers in the population that were either clear record or accident drivers. This means that approximately 40 future good drivers would receive failure scores for every one future bad driver receiving a failure score. The implementation of such a system is likely to be impractical.

The discussion so far has been on the pessimestic side, which perhaps is misleading. The primary results of the study show that various categories of drivers can successfully be distinguished on the basis of patterns of scores on driver skill tests. Reports in the literature so far have pointed to individual driving skills (Campbell, 1958) as possible distinguishers, or have indicated non-significant results, but none has reported patterns of skills that yield significant differences between categories of drivers. These patterns of skills were given in Table 7. A look at this table indicates that two different classes of skills seem to contribute: First, a "physical handling of the automobile" class, including

brake stop, turn about, stop and start, and clutch. The second class is an "interaction with traffic" class, including attention, keeping in lane, right of way, and first slow sign. With the identification of patterns of skills, further research may now be designed to investigate these patterns, perhaps identifying an underlying complex of skills that yields better results than the systems identified in this study.

Finally, it should be pointed out that a number of other types of variables have also been shown to discriminate significantly between drivers: for instance, measures of driver attitudes, previous driver performance, etc. This discrimination has generally been found using statistical methods somewhat less powerful than the one used in this study. The multivariate analysis of variance paradigm could be applied to these areas of research, perhaps yielding clearer results than now are available. In addition, the multivariate analysis of variance paradigm could be used to study the interrelation of all the classes of variables in an attempt to get a general overview to the relevant variables that might be used to successfully distinguish between categories of drivers.

Conclusions. The research reported in this paper is only a modest beginning. The results of this research provide no clear answer to the question "Which driving skills are most important?" but they do offer some insight into the patterns of skills most important. The patterns identified may now be further studied to more sharply define them and to get a closer approximation to their predictive ability.

A number of suggestions concerning future research were given above. Of these, the exposure to driving variable is the most important to control in future research. The composition of the categories of drivers defined for this study is still an unknown due to the lack of control over the exposure variable.

In conclusion, it can be said that this study indicates a future for the "driver skills" class of variables in the differentiation and prediction of future driver performance. The results are not conclusive, but do seem to be consistent. This class of variables may not lead to adequate prediction of driver performance by itself, but it is possible that in conjunction with other classes of variables, better prediction can be made.

References

American Medical Association. The state of the art in traffic safety. Report prepared by Arthur D. Little, Inc., Acorn Park, Cambridge, Massachusetts, June, 1966.

Bock, R.D. Estimating multinomial response relations. Research Memorandum Number 5, Statistical Laboratory, Department of Education, University of Chicago, Chicago, Illinois. July, 1967.

Campbell, B.J. A comparison of the driving records of 1100 operators involved in fatal accidents and 1100 operators selected at random. Traffic Safety Research Review, 1958, 2(3), 2-7.

Clyde, D.J., Cramer, E.M., and Sherin, R.J. Multivariate Statistical Programs. Coral Gables, Florida: Biometric Laboratory, University of Miami, 1966.

Goldstein, L.G. Research on human variables in safe motor vehicle operation: A correlational summary of predictor variables and criterion measures. The Driver Behavior Research Project, The George Washington University, Washington, D.C., 1961.

Goldstein, L.G. Human variables in traffic accidents: A Digest of research. Traffic Safety Research Review, 1964, 8(1), 26-31.

Jones, L.V. Analysis of variance in its multivariate developments. In Handbook of Multivariate Experimental Psychology (Raymond B. Cattell, Editor), Rand McNally and Company, Chicago, Illinois. 1966.

Levonian, E. Prediction of accidents and convictions. Traffic Safety Research Review, 1967, 10(3), 75-79.

Rao, C.R. Linear Statistical Inference and Its Applications. John Wiley and Sons, Inc. New York, New York. 1965.

Schuster, D.H. Long term trends in automobile accident research. Paper presented at a meeting of the National Independent Statistical Service, September, 1966.

Schuster, D.H. Prediction of current accident records for youthful drivers. Traffic Safety Research Review. 1966, 10(1), 22-25.

Uhlaner, J.E. Selection tests--dubious aid in driver licensing. Communication from the National Safety Council, 425 N. Michigan Avenue, Chicago, Illinois. 1966.