The development and evaluation of the North Carolina pictorial oral driver license examinations were supported in part by the Governor's Highway Safety Program.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

The development of the new pictorial oral driver license examinations would not have been possible without the dedication and effort of a great number of people. It would not be possible to mention everyone by name but there are some whose contributions were particularly significant.

The art work was provided by Rebecca McCollum and Lauren Ogle, with William Pope providing photographic assistance.

The North Carolina Division of Motor Vehicles was prominently involved from the very beginning, and their help and support were indispensable. Joseph Price provided valuable criticism of the items. J. William Lackey, Lorraine S. de Savornin Lohman, and Susan S. Padgett participated in the training of the driver license examiners. Sandra Hough provided a variety of skills throughout the project.

The input provided by the driver license examiners led to revisions in virtually every item. Their thoughtful and considered criticism is reflected in the final results. The examiners also provided the data that were used for the evaluation, since, of course, the applicants involved were not able to fill out data forms. The examiners also provided considerable moral support, much needed for a project that extended over several years.

ABSTRACT

A need was identified for developing new driver license knowledge tests for use with applicants unable to read well enough to take the written version of the test. The new pictorial oral exams were based to a large extent on the efforts that had previously gone into the development of new written tests for North Carolina. However, each item was presented pictorially with the driver license examiner explaining the question and the answer choices to the applicants. There was considerable input from driver license examiners as well as others in the development of the pictorial items.

The new tests were introduced statewide in December of 1974, and special data were collected from the field in January of 1975. Information on renewal applicants was linked to their prior driver records, and a variety of psychometric analyses were performed.

The results of the analyses included the following:

1. Applicants taking the oral version of the test differ markedly from those taking the written form. The oral exam applicants are, on the whole, older, much more likely to be male, much less likely to be white, and not as well educated as written exam applicants.

2. The mean scores on the three versions of the test were fairly close, averaging about 17 correct items. Nevertheless, a passing score required at least 18 correct items. It should be noted, however, that the mean scores on the new written tests were about 16 items correct, indicating that the new pictorial oral tests are slightly less difficult than the written rules examinations.

3. The test reliabilities are all at an acceptable level and are very close to those reported for the new written tests.

4. On the whole the item means indicate that the test items are at an acceptable level of difficulty.

5. The correlations between items and total test performance were generally acceptably high, while the inter-item correlations were generally low, indicating that on the whole the tests are assessing one general set of knowledge but from a variety of independent aspects.

6. The relationship between test performance and driver record variables appears stronger for reckless violations and for accidents than for other violation types or total violations.

7. Items were classified according to three categories, namely, situational questions dealing with predicaments in which a driver may find himself, procedural questions that require the applicant to describe safe driving procedures, and administrative rules and general rules of good driving. It was found that the situational questions were more predictive of the driver record variables than were the other two types of questions.

8. Questions were finally considered on the basis of five criteria: difficulty level, item-total test correlation, number of times a question was used in predicting previous driving record, item's predictive power regarding accidents, and the direction of the relationship between item performance and driver record variables. The results of these evaluations led to the final selection of items.

9. Two final versions of the test are recommended, Form Alpha and Form Beta. These two tests have been equated for predicted overall difficulty level and for predicted item-total test performance. In addition, they each cover the same number of situational, procedural, and administrative questions. The order of items is the same for the different types of questions except for minor variations that were allowed in order to place unusually easy items at strategic points. Thus, for each version of the test the first, third, and final items are particularly easy. The two recommended versions of the test should provide two parallel driver license examinations that have meaningful relationships to driver records of applicants who do not read well enough to take the written tests. Ideally, these two new test forms should be field tested to insure that they indeed perform as well as would be anticipated on the basis of the present analysis.

INTRODUCTION

North Carolina has been engaged in a long range project concerned with evaluating and upgrading the driver licensing program. As part of this effort, new written knowledge examinations have been developed specifically based on North Carolina drivers. These tests were introduced statewide in December of 1972.

A significant portion of North Carolina driver license applicants do not take the written version of the rules examination because they have difficulty in reading. Because reading ability does not necessarily correspond to good driving practices, an oral knowledge test has been administered to these applicants. The oral exam was offered in an attempt to measure driving knowledge independently of reading skill. The oral examination that was used prior to December, 1974 consisted of open-ended questions, and each examiner had available to him an answer sheet providing acceptable responses.

This oral examination, however, was not a parallel form of the written examination. That is to say, the oral examination differed from the written examination in two critical features. First, the written examinations provided four alternative responses to each question so that the applicant had only to recognize the correct response from among several possibilities. In addition, this format meant that even if the applicant taking the written form of the test had no knowledge about the question, he had a one in four chance of guessing the correct answer. The oral examination, on the other hand, required that the applicant recall the correct response and formulate a correct answer verbally. This recall and verbalization task is much more difficult than simply recognizing the correct response. The verbalization is doubly difficult for the applicant since persons taking the oral exam are probably less able in terms of verbal skills than applicants taking the written rules examination. The oral examination format did not allow for guessing correctly which made the exam more difficult.

The second difference between the two types of examinations deals with the form of the correct answer to a question. In the written exam the correct answer was obvious, easy to score and open to no interpretation by the examiner. In the oral test, however, the examiner had to interpret the answer given by the applicant. This led to difficulties in grading: paraphrasing research has shown that people have difficulty in correctly identifying true paraphrases and distinguishing false paraphrases from true paraphrases. The open-ended format might also allow for examiner biases (either conscious or unconscious) to interfere with the testing situation.

In the process of developing the new written tests a decision was made to consider developing better tests for the oral exam applicant. Because at that time state summary statistics did not differentiate between oral and written tests, a survey was conducted to ascertain how many applicants took the oral version of the examination (McMichael and Waller, 1973). A sample of driver license examiners was selected to include four examiners from each of the eight driver license districts. Within each district two of the examiners were from urban stations and two from rural stations. These 32 examiners were sent data sheets and asked to compile information for a one-week period. The information requested included the number of tests given by type of test (oral or written), type of applicant (learner's permit, original, or renewal), and the outcome of the test (pass or fail). It was found that permit applicants were somewhat less likely to take the oral exam than original and renewal applicants. Persons applying at rural stations were somewhat more likely to take the oral test than those coming to urban stations. About 12 percent of all applicants in the sample took the oral tests, indicating that the oral exam is a significant part of the driver licensing program. In a year's time this would represent over 100,000 applicants.

Based on the survey results and on the obvious shortcomings of the existing oral exams, a decision was made to proceed with the development of new instruments.

This report describes the development of the oral exam items and tests, the collection of performance data, the evaluation of the tests based on the performance of applicants in the field, and recommendations for the final composition of the oral exams.

METHOD

The new tests attempted to present questions that were easier for the applicant to understand and to respond to and easier for the examiner to score. Each question was to be presented along with pictures that would represent four different alternatives that the applicant could choose in making his response. To a large extent the new tests were built on the effort that had already gone into the development of the new written tests. In many instances written items were adapted to pictorial presentation. Other new items were developed for use on the oral exams.

Initially an effort was made to get color photographs illustrating the questions. Two problems quickly became evident. Photographs often contained distracting factors that had nothing to do with the question. The other problem was that it simply was not possible to get photographs of the great variety of situations and conditions needed for the questions. Therefore, a decision was made to employ the skills of a graphic artist.

Items were outlined in cooperation with the Division of Motor Vehicles (DMV) personnel. As they were completed they were reviewed by additional DMV personnel for accuracy. Specific questions about signs or road markings were checked out with the Highway Commission and questions of a legal nature were clarified with the help of the University of North Carolina Institute of Government. When the first set of items was completed it was presented at the annual in-service training school for driver license examiners in the fall of 1973. Each examiner was given a set of Xerox copies of the items, and items were reviewed using projected color slides of the original pictures. Scarcely an item survived this review process without some revision, and some items were eliminated entirely. Five months later the next two sets of items were ready for review. A special group of 20 examiners was selected from throughout the state. These examiners were chosen because they had extensive experience working with applicants who had reading difficulty. Again each examiner had Xerox copies of the items and in addition the items were presented in color slides. A full day was spent reviewing these two sets of items. Once again, very few items survived without some modification, and again some items were eliminated entirely.

Some of the examiners participating in this special meeting indicated that they worked at stations that not only had a fair number of applicants taking the oral exam but also had physical facilities such

that an oral exam could be presented in slide form without disturbing other applicants. Six examining stations located throughout the state were visited, and the test items were presented in slide format. The slides were used as the oral exam for applicants who could not read well enough to take the written test. The information gained from these experiences was also used in making final revisions in the test items.

In developing the oral tests it was necessary to anticipate as many problems as possible ahead of time. Unlike the written test, where it was fairly easy to field test items and make revisions, the pictorial oral exam would not easily lend itself to revisions once the pictures were printed. Thus it was essential that we engage in the time-consuming and meticulous process of obtaining detailed input from the driver license examiners on test items before they were prepared in final form.

Several points should be made about the tests that were finally prepared. First, the items were based on areas of knowledge that were considered important by DMV personnel. Second, each of the original three versions of the test covered the same areas of content and to the same extent as the other forms. Third, all questions were reviewed by DMV personnel and checked for legal accuracy. Fourth, every item was reviewed in detail by at least 20 experienced and qualified driver license examiners. Fifth, the pictorial exams were not intended to be easier than the old oral exams. The intent was to hold the non-reading applicant responsible for mastering the same body of information required from literate applicants -- to create a pictorial form of the rules test that was parallel in form to the written tests. The new tests, however, were designed to make it easier for the applicant to understand what is expected of him and to communicate to the examiner how much he knows. In addition, the examiner should find the new oral tests easier to score than the old ones.

In November of 1974 a series of one-day workshops was conducted throughout North Carolina, one in each of the eight driver licensing districts. At these workshops each examiner was given his own set of tests and instructed how to administer each item. The test is administered individually, and the examiner emphasizes relevant portions of the picture choices as he reviews them with the applicant. Each test is preceded by a sample item that is used to acquaint the applicant with how to respond to the new test. Appendix A gives the instructions to the examiners on how to administer the tests.

In early December of 1974 the new oral examinations went into use statewide and have since been the official oral driver license examination in North Carolina. In January of 1975 the driver license examiners collected special data from applicants taking the oral tests and provided

detailed information on how the applicants performed on each test item. Psychometric analyses were performed based on renewal applicants. The results were used as the basis for making final revisions in the tests.

Only data from renewal applicants are analyzed in this paper. The decision to look at renewals rather than original applicants was made so that it would be possible to explore the actual relation between the new oral examinations and driving behavior without waiting several years. By looking at an applicant's previous driving record, one can estimate how good a driver the applicant is and then examine the relationship between the test results and the actual driving behavior.

RESULTS AND DISCUSSION

Description of Population

Table 1 shows data based on renewal applicants taking the oral exam and renewal applicants taking the written version of the test. The data on the oral exam applicants were collected in January of 1974, while the data on the written tests were collected during January of 1973.

It can be seen that the two groups of applicants are quite different. The oral exam applicant is older (in his late 40's rather than his mid to late 30's). He is much more likely to be male (94 percent compared to only 53 percent for the written test applicant) and much less likely to be white (37 percent compared to 86 percent for the written test applicant). Finally, as would be anticipated, the applicant taking the oral exam is not as well educated as those taking the written tests. Indeed, over three-fourths (76.3 percent) of the oral exam applicants had only an elementary education or less. The applicant taking the written test was likely to have completed high school. (It should be noted that many renewal applicants undoubtedly attended school in North Carolina at a time that a high school diploma was awarded after completing eleven grades. Thus it is likely that the obtained figures represent, on the whole, about the equivalent of a high school diploma.)

Test Analysis

The first analysis that was performed involved a standard test-item analysis of the three different forms of the new oral rules test (Magnusson, 1968). This analysis yielded response distributions, alpha and Spearman-Brown reliability coefficients, test means, item means, and item-total correlations for each of the test forms.

Table 1. Comparison of applicants taking the oral and written versions of the driver license knowledge examination.

Applicant Characteristic	Oral	Written
Age (in years)		
Mean	48.3	Males - 39[a] Females - 37[a]
Median	48.1	Males - 36[a] Females - 35[a]
Sex (percentage)		
Male	93.9	53.3[a]
Female	6.1	46.7[a]
Race (percentage)		
White	37.3	85.7%[b]
Non-white	62.7	14.3%[b]
Education (in years)		
Mean	4.5	Male 12.3[a] Female - 12.3[a]
Median	4.0	Male - 11.7[a] Female - 11.9[a]

[a] From Creech and Grandy, 1974

[b] From Stewart, 1975

Tables 2 through 4 indicate the distribution of the responses for each item and each test form. The correct response for each item is underlined. The last two columns show the number of multiple responses and the number of omitted responses. No item seems to be predominately inappropriately responded to (either two or no responses). This indicates that the applicants seem to have a good understanding of the test procedure. The question that was most often not responded to (or was responded to with multiple answers) was question A18 (question 18 on form A) with 11 out of 232 (4.7 percent) responses in this category. The different test forms showed a similar pattern of two or more responses (0.7 percent, 0.9 percent, and 0.6 percent, respectively). Test form C, however, seemed to elicit fewer no responses (1.4 percent, 1.7 percent, and 0.6 percent, respectively).

Ideally one would want questions where the correct answer is chosen most often while the three incorrect alternatives are about equal in the number of responses they elicit. In examining those questions in which the correct answer was given half or less than half the time (question A1, A3, A5, A12, A18, A20, B11, B17, and C4), we find two different patterns of responses. In one pattern the two most frequent responses clearly dominate the other alternatives. Questions A1, A5, and A18 are of this type. Clearly in these questions the third and fourth alternatives are not seen by the applicants as being true alternatives. There might also seem to be a certain amount of ambiguity between the two dominant alternatives. If these questions are to be retained, it is suggested that new alternatives be substituted and the two prime alternatives be examined for clarity.

The second pattern of responses, which includes the other questions listed above, does not indicate the previous response bias. Rather, these questions show the expected slower decrease in the frequency of selection. It is suggested that these items might be either ambiguous or difficult. Which case is correct can be determined by attempting to rephrase the question in a more straight-forward manner. If the question still shows this particular pattern of responses, then the question is dealing with an area of knowledge which is apparently misunderstood by the population and needs to be emphasized in the driver manual.

The test means appear particularly low (see Tables 5 through 7). Applicants must get 18 items correct in order to pass the test, and all three test forms have means lower than 18. Figure 1 indicates the distribution of test scores for the three different test forms.

The poor performance on the tests can be interpreted in terms of the population characteristics of the applicants taking the oral renewal

Table 2. Response distributions for test form A.

Question #	A	B	C	D	2 Responses	No Response
1	82	20	28	99	2	1
2	13	24	163	25	2	5
3	37	28	116	45	1	5
4	16	28	169	19	0	0
5	25	76	19	105	6	1
6	29	5	33	158	1	6
7	154	36	14	22	1	5
8	134	42	11	40	1	4
9	13	37	12	166	0	5
10	201	9	3	16	0	3
11	16	201	4	7	1	3
12	53	74	23	75	3	4
13	15	22	162	27	3	3
14	17	4	133	73	2	3
15	39	10	150	29	1	3
16	12	25	121	72	0	2
17	12	14	17	186	0	3
18	6	109	26	80	5	6
19	14	7	193	14	1	3
20	58	39	93	37	2	3
21	17	38	14	158	2	3
22	12	19	166	30	1	4
23	17	180	6	24	2	3
24	11	27	12	178	1	3
25	25	10	7	187	0	3
Total					38	84
Percent					.007	.014

Table 3. Response distributions for test form B.

Question #	A	B	C	D	2 Responses	No Response
1	73	9	8	50	2	3
2	2	9	127	4	1	2
3	11	108	12	10	1	3
4	17	39	78	5	2	4
5	5	120	2	16	1	1
6	127	4	3	8	0	3
7	8	86	3	47	0	1
8	12	87	14	29	2	1
9	27	74	23	15	2	4
10	10	15	14	104	1	1
11	18	64	34	23	3	3
12	108	5	6	23	1	2
13	11	3	119	10	0	2
14	105	11	8	17	2	2
15	9	105	12	15	2	2
16	2	14	3	123	0	3
17	21	15	64	39	3	3
18	3	4	8	129	0	1
19	2	9	121	8	1	4
20	14	101	9	20	0	1
21	10	124	9	1	0	1
22	14	78	15	31	3	4
23	4	36	14	85	3	3
24	9	89	27	14	2	4
25	36	91	6	10	0	2
Total					32	60
Percent					.009	.017

Table 4. Response distributions for test form C.

Question #	A	B	C	D	2 Responses	No Response
1	9	19	60	24	1	0
2	7	61	27	16	1	1
3	9	86	7	11	0	0
4	34	42	21	14	0	2
5	69	20	19	5	0	0
6	6	82	7	17	0	1
7	32	16	3	57	2	3
8	19	4	2	87	0	1
9	83	12	9	8	1	0
10	12	60	4	34	2	1
11	2	0	100	11	0	0
12	12	7	77	15	1	1
13	7	6	23	75	1	1
14	104	3	3	1	2	0
15	16	78	8	10	1	0
16	8	5	7	92	1	0
17	65	4	26	17	1	0
18	5	3	99	5	0	1
19	11	74	18	9	1	0
20	106	2	2	2	1	0
21	90	5	6	11	0	1
22	2	7	5	99	0	0
23	4	77	18	13	0	1
24	6	91	6	8	1	1
25	8	2	90	11	1	1
Total					18	16
Percent					.006	.006

Table 5. Test analyses of form A.

Number of Applicants	232
Test Mean	16.065
Test Variance	21.707
Reliability Estimates:	
Alpha = 0.764	
Spearman Brown Correction =	0.769

Mean and Item-Total Correlations

Item	Mean	Correlation	Item	Mean	Correlation
1	0.427	0.270	14	0.573	0.472
2	0.703	0.515	15	0.647	0.355
3	0.500	0.484	16	0.522	0.528
4	0.728	0.350	17	0.802	0.441
5	0.453	0.439	18	0.345	0.272
6	0.681	0.343	19	0.832	0.565
7	0.664	0.451	20	0.401	0.395
8	0.578	0.446	21	0.681	0.484
9	0.716	0.288	22	0.716	0.505
10	0.866	0.416	23	0.776	0.471
11	0.866	0.356	24	0.767	0.413
12	0.319	0.324	25	0.806	0.461
13	0.698	0.336			

Table 6. Test analyses of form B.

Number of Applicants	145
Test Mean	17.172
Test Variance	21.481

Reliability Estimates:
Alpha = 0.773
Spearman Brown Correction = 0.846

Means and Item-Total Correlations

Item	Mean	Correlation	Item	Mean	Correlation
1	0.503	0.406	14	0.724	0.343
2	0.876	0.375	15	0.724	0.412
3	0.745	0.483	16	0.848	0.364
4	0.538	0.408	17	0.441	0.470
5	0.828	0.565	18	0.890	0.460
6	0.876	0.452	19	0.834	0.409
7	0.593	0.440	20	0.697	0.419
8	0.600	0.489	21	0.855	0.421
9	0.510	0.432	22	0.538	0.372
10	0.717	0.456	23	0.586	0.345
11	0.441	0.521	24	0.614	0.283
12	0.745	0.513	25	0.628	0.450
13	0.821	0.320			

Table 7. Test analyses of form C.

Number of Applicants	113
Test Mean	17.735
Test Variance	18.797

Reliability Estimates:
Alpha = 0.753
Spearman Brown Correction = 0.722

Means and Item-Total Correlations

Item	Mean	Correlation	Item	Mean	Correlation
1	0.531	0.417	14	0.920	0.314
2	0.540	0.419	15	0.690	0.524
3	0.761	0.516	16	0.814	0.475
4	0.372	0.372	17	0.575	0.476
5	0.611	0.160	18	0.876	0.293
6	0.726	0.525	19	0.655	0.509
7	0.504	0.319	20	0.938	0.272
8	0.770	0.466	21	0.796	0.223
9	0.735	0.402	22	0.876	0.336
10	0.531	0.331	23	0.681	0.488
11	0.885	0.330	24	0.805	0.552
12	0.681	0.501	25	0.796	0.435
13	0.664	0.414			

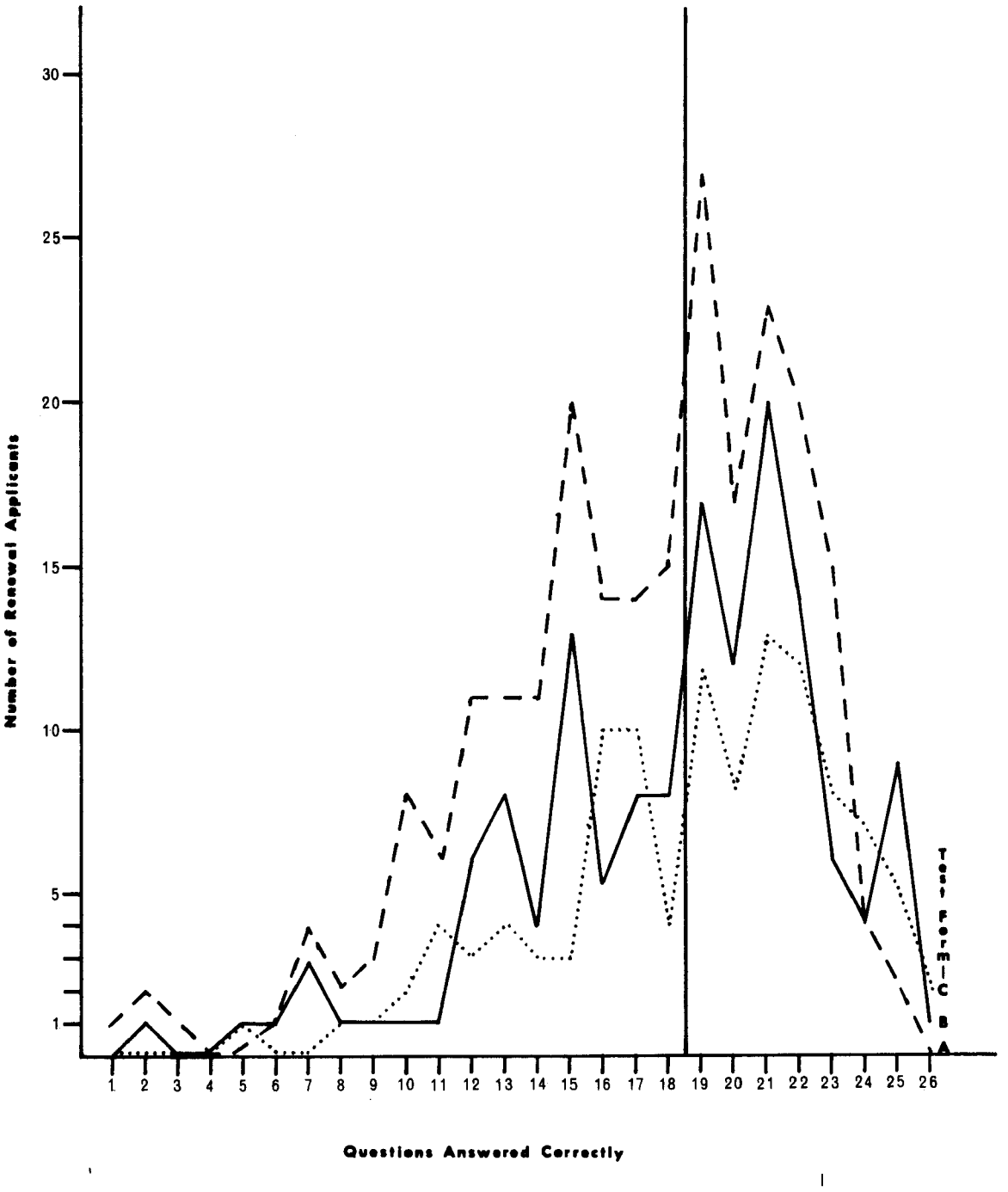

Figure 1. Total score distributions for each test form.

examination. Since the population is characterized by a low educational level, it would have been surprising if the applicants had performed at a passing level. The consistence of the test means would seem to indicate that the low scores are a function of the population characteristics rather than the test characteristics.

On the other hand, Creech and Grandy (1974) report that applicants taking the renewal written rules tests had a mean score of 15.8 correct answers. This means that the new pictorial oral tests are slightly less difficult than the written rules test.

The test reliabilities (Tables 5, 6, and 7) indicate that the tests seem to be consistently assessing the same general concepts. The alpha coefficient represents the lower limit of the reliability of a test. The Spearman-Brown coefficient is the more frequently reported figure. These coefficients seem to indicate that the new oral exams are as reliable as the written exams as analyzed by Creech and Grandy, who reported reliabilities that were approximately equal to these. Test B appears to be the most reliable of the tests. The over-estimation of alpha in test form C (i.e., alpha is larger than the Spearman-Brown coefficient) is probably due to the small sample size.

The item means (Tables 5, 6, and 7) represent the best description of the difficulty level of the test items. As a general rule, one prefers not to have items which are either too easy or too difficult. For example, questions A5, A12, A18, A20, B11, B17, and C4 are each missed by 55 percent to 70 percent of the applicants. It may be the case that these items are ambiguous and need to be rewritten. Items which are too easy should usually be avoided, also. Questions B2, B6, B18, C11, C14, C18, C20 and C22 are all successfully answered by at least 87.5 percent of the applicants. While these questions probably have some value in terms of relaxing the applicants and reducing the stress of the situation, the questions could be revised to provide more viable alternatives.

The item-total correlations (the correlation between performance on the item and the total test score) are useful in identifying items which appear to be assessing different concepts than the test as a whole. Items with especially low item-total correlations (A1, A18, B24, C4, C20, and C21) should be examined for relevance to the concepts tested in the exams.

Tables B-1, B-2, and B-3 in Appendix B contain a correlation matrix for each test. Note that the inter-item correlations (the correlations between performances on each item and performances on every other item) are relatively low, indicating that the questions are not repetitive in

nature. Also, there are few negative correlations indicating that the items appear to be of a single nature with regard to measuring overall driving knowledge. The tests on the whole seem to be examining one central theme but approaching it from a number of independent aspects.

The correlations were submitted to a principal components analysis (Tatsuoka, 1971), and one general factor was found to be accounting for 18.1 percent, 18.6 percent, and 17.6 percent of the variance of each test, respectively. This is the usual finding for tests of this nature. The factor loadings (not reported) were not clearly interpretable.

Relation to Previous Driver Record

Since driving records were available over a four-year period previous to the exam, attempts were made to predict eight prior driving variables for each year from the questions on the different test forms. The driving record variables that were explored were number of speeding violations, stop violations, moving violations, reckless violations, total number of violations, number of accidents at fault, number of accidents not at fault, and the total number of accidents.

A stepwise regression analysis was performed for each driving record variable for each time period with a cutoff F level of 3.0. The results of these analyses are presented in Tables 8, 9, and 10. The first column indicates the question number. The regression coefficient in the final regression line is listed in the second column. The F level at which each variable is entered is indicated in the third column. The last column records the multiple correlation at each level of the stepwise regression. These final multiple correlations range from .12 to .60, accounting for up to 36 percent of the variance of the previous driving record.

Table 11 indicates the mean multiple correlations for various subgroups. Three of these results are of particular interest. First, note that test form C is much more predictive of previous driving behavior than test forms A and B (.31>.23, .31>.19). Furthermore, test form B appears to be more predictive than test form A, but this effect is less marked.

Second, the predictive power of the tests varies over the years. (see Table 12). The second year prior to the renewal test is by far the best in terms of prediction. The lower correlations for the first year prior to renewal require some further explanation. Violations are much more frequent events than accidents. However, a violation is not entered on the driver record unless it results in a conviction, and

FORM A

Table 8. Summary regression table for test form A.

Driving Record Var	Question No.	Regression Coefficient	F Level Entered	Mult. Correlation	Question No.	Regression Coefficient	F Level Entered	Mult. Correlation	Question No.	Regression Coefficient	F Level Entered	Mult. Correlation	Question No.	Regression Coefficient	F Level Entered	Mult. Correlation
Speeding	3	.14	4.6	.14	13	.14	4.8	.14	14	.16	4.5	.14	14	.16	8.0	.18
	20	-.13	3.9	.19					13	-.16	6.0	.21	13	.14	4.7	.23
	8	.12	3.1	.22												
	24	-.11	3.0	.25												
Stop	15	-.13	3.6	.12	2	.22	5.9	.16	19	-.16	12.5	.23	14	-.19	7.4	.18
	3	.15	4.4	.18	21	-.22	8.5	.25	7	-.15	5.1	.27	13	.13	3.9	.22
	12	-.13	3.7	.22	10	.15	4.3	.28	1	.17	6.1	.31				
					15	-.12	3.7	.30	11	-.13	3.9	.33				
					5	-.13	3.0	.32	10	-.12	3.1	.35				
					13	.12	3.7	.35	13	.11	3.0	.37				
Moving					23	.13	3.0	.11	2	-.19	6.5	.17	13	.12	4.2	.13
					15	-.13	3.6	.17	16	.13	4.0	.21	4	.11	3.1	.18
Reckless	10	-.15	4.2	.13	22	-.20	7.5	.18	24	-.17	6.0	.16	7	.14	4.7	.14
	12	.14	4.6	.19	20	.14	4.8	.23	12	.11	3.1	.20	20	.11	3.4	.19
Accident At Fault	14	-.13	4.1	.13	12	-.20	8.8	.19	12	.12	4.1	.13	13	.14	4.6	.14
					20	.12	3.7	.23	9	-.15	3.9	.18				
									23	.11	3.1	.22				
Total Violations	3	.14	5.0	.15	19	.16	4.0	.13	24	-.17	7.1	.17	21	.12	3.4	.12
	10	-.15	4.5	.20	21	-.12	3.3	.18	12	.15	4.4	.22				
	23	.15	4.7	.24					23	-.12	3.2	.25				
	20	-.12	3.4	.27												
Total Accidents	10	-.19	5.0	.15	24	.14	3.1	.12	12	.16	6.3	.16				
	21	.15	4.1	.20	21	-.14	4.1	.18								
	14	-.15	3.3	.23												
	7	.12	3.3	.26												
Accident Not Fault	10	-.17	4.8	.14	12	.13	4.5	.14					9	-.14	4.6	.14
	2	.13	5.5	.21	24	.12	3.3	.18								
	5	.12	3.3	.24												

FORM B

Table 9. Summary regression table for test form B.

Driving Record Var	Question No.	Regression Coefficient	F Level Entered	Mult. Correlation	Question No.	Regression Coefficient	F Level Entered	Mult. Correlation	Question No.	Regression Coefficient	F Level Entered	Mult. Correlation	Question No.	Regression Coefficient	F Level Entered	Mult. Correlation
Speeding					21	-.18	4.6	.18	11	.17	7.6	.22				
					24	.15	3.5	.23	10	.16	3.4	.27				
									24	.14	3.0	.30				
Stop	20	.16	3.2	.15	7	-.22	7.2	.22	24	-.19	3.2	.15				
	8	-.19	3.6	.22	22	.19	5.3	.29	16	.16	3.4	.21				
	25	.15	3.0	.26	11	-.19	4.5	.33								
					24	.18	4.9	.38								
Moving					11	.21	4.1	.24	25	-.16	3.6	.16	21	-.18	4.9	.18
					2	-.18	4.3	.29	9	.20	4.3	.23				
					8	-.18	3.2	.32	8	-.15	3.1	.27				
Reckless	13	-.18	4.7	.18	10	.22	5.4	.19	1	-.18	3.2	.15	9	.25	7.9	.23
					25	-.23	6.7	.28	8	.17	4.2	.22	23	-.21	5.4	.28
					8	-.19	3.1	.32	18	-.19	3.7	.27	7	-.21	4.9	.34
					5	.17	4.0	.35	23	.16	4.0	.32	12	.16	3.8	.38
					1	.18	3.4	.38								
					24	-.15	3.3	.41								
Accident At Fault					15	-.31	9.3	.25	1	-.16	3.4	.15	9	.16	3.6	.16
					1	.16	3.5	.29	5	.21	3.7	.22	13	-.20	4.3	.23
					14	.15	3.3	.32	25	-.17	4.1	.27	17	-.20	3.4	.28
													16	.16	3.9	.32
													22	.15	3.3	.35
Total Violations	25	.16	3.6	.16	10	.23	5.4	.19	23	.15	3.2	.15				
					23	-.24	7.7	.29								
					24	.16	4.2	.34								
Total Accidents	9	.16	3.6	.16	4	.17	4.8	.18	1	-.26	6.8	.21	13	-.27	9.9	.25
					1	.20	3.8	.24	11	.20	5.6	.29	22	.23	7.4	.33
					15	-.19	5.5	.31					16	.18	4.1	.37
													17	-.17	4.7	.41
													7	-.17	3.1	.43
													9	.15	3.7	.45
Accident Not Fault	9	.20	6.2	.20	4	.26	8.6	.24	1	-.20	3.6	.16	13	-.18	7.5	.22
					1	.18	3.3	.28	11	.20	5.6	.25	25	.19	5.9	.29
					11	-.16	3.8	.32					7	-.21	5.2	.34
													22	.16	4.2	.38

FORM C

Table 10. Summary regression table for test form C.

Driving Record Var	Question No.	Regression Coefficient	F Level Entered	Mult. Correlation	Question No.	Regression Coefficient	F Level Entered	Mult. Correlation	Question No.	Regression Coefficient	F Level Entered	Mult. Correlation	Question No.	Regression Coefficient	F Level Entered	Mult. Correlation
Speeding	6	.18	4.85	.20	25	.16	3.1	.16	17	.17	5.3	.21	19	.18	3.5	.18
	4	-.26	6.2	.30					15	.17	3.2	.27				
	9	.20	4.4	.36												
Stop	13	-.24	8.7	.27	22	-.26	14.6	.34	6	-.21	3.3	.17	6	-.23	5.0	.21
	21	-.23	8.4	.37	3	-.24	6.4	.41	1	.20	4.8	.26	24	.18	3.1	.26
	11	-.16	3.0	.40	17	.19	3.6	.44					7	-.16	3.1	.31
					4	-.18	4.2	.47								
Moving	23	-.18	3.6	.18	3	-.31	17.9	.37	14	-.30	5.7	.22	25	-.20	4.9	.20
					22	-.29	10.2	.46	23	.24	5.7	.31				
					14	.19	4.2	.49	17	-.21	3.4	.35				
					13	-.17	4.2	.52	22	-.20	3.2	.39				
									2	.19	3.3	.42				
									11	.18	4.1	.45				
Reckless	20	-.24	7.0	.24	16	-.24	4.2	.19	11	-.32	8.4	.26	25	-.26	8.9	.27
					15	.25	6.5	.30	16	.30	7.4	.36	18	-.24	6.9	.36
					11	-.19	3.1	.34	17	.27	4.9	.41	16	.18	4.0	.40
					9	.16	3.1	.38	15	-.21	5.3	.45				
									18	-.20	4.2	.48				
									5	-.16	3.7	.51				
Accident At Fault	2	-.21	5.0	.21	22	-.35	15.4	.35	2	.23	4.8	.20	11	-.29	7.5	.25
					13	-.23	7.2	.42	6	-.21	5.1	.29	23	-.22	4.2	.31
													3	.18	3.6	.36
Total Violations	6	.20	4.5	.20	22	-.28	15.9	.35	17	.19	4.3	.19				
					3	-.25	5.6	.41								
					17	.20	5.2	.45								
Total Accidents	2	-.20	4.6	.20	22	-.40	24.0	.42	4	.22	4.3	.19	11	-.21	5.4	.21
					19	-.27	7.9	.48	12	-.21	3.2	.25				
					15	.23	4.0	.51	16	.20	4.7	.32				
					3	-.22	7.0	.55								
					13	-.18	3.0	.57								
					20	.15	3.2	.59								
					14	.15	3.2	.60								
Accident Not Fault	5	-.20	4.8	.20	22	-.22	5.5	.22	12	-.29	5.8	.22	6	-.26	5.1	.21
					6	.20	3.3	.28	3	.19	5.9	.31	10	.18	3.5	.27
					19	-.18	3.7	.33	17	.19	4.2	.36				

Table 11. Mean multiple correlations for previous driving record.

	Test Form			
	A	B	C	
Speed	.21	.13	.24	.19
Stop	.29	.16	.36	.27
Moving	.14	.19	.31	.21
Reckless	.20	.32	.38	.30
Total AccFault	.18	.24	.32	.25
Total Viol	.21	.16	.21	.19
Total Acc	.15	.30	.33	.26
Total AccNotFault	.14	.29	.29	.24
	.19	.23	.31	.25

Table 12. Mean multiple correlations over previous four years for each test form.

	First Year Prior	Second Year Prior	Third Year Prior	Fourth Year Prior
A	.20	.21	.20	.17
B	.12	.33	.26	.22
C	.25	.42	.33	.24
	.19	.32	.26	.21

the processing through the court system often takes several months. The copy of the driver file that was used in this evaluation was made in early April, meaning that there were undoubtedly some violations incurred by these January applicants prior to their renewal that had not yet been entered on the file. Thus the driver record data for the first year prior to renewal could not be considered as complete as those for the other three years. Whether on the basis of more complete data the correlations for this first prior year would reach the level of the second and third years prior to renewal cannot be determined on the basis of these data. However, the reasonably high correlations for years two and three indicate that the test seems to be measuring general driving behavior and not simply behavior that may be the result of recent learning of the rules of the road.

Lastly, note that the tests are more predictive of accidents and reckless violations than of total violations (.26>.19, and .30>.19, see Table 14).

Returning to Tables 8, 9, and 10, it is also possible to interpret the signs of the regression coefficients. A negative coefficient would infer that either someone answered a question incorrectly and was previously involved in a violation or accident, or someone answered the question correctly and was not involved in a violation or accident. This is the direction of relationship that is desired. Some questions, such as A10, A15, B7, B13, C22, and C6, seem to predict predominately in the correct direction. Others, such as A12, A13, B9, B22, C16, and C17, predict in an incorrect fashion. This particular aspect of the questions would appear to be critical in determining what questions should be retained. Table 13 presents a simple numerical count of the number of times each question had a positive and a negative coefficient.

Tables 14, 15, and 16 reproduce the information in Tables 8, 9, an 10 organized by question. They provide a short description of the content of each question, which driving record variable(s) it is used in predicting, and finally, in parentheses, for which time period it provides prediction.

Table 17 categorizes the different questions into different content areas. In the initial development of the oral exams, an attempt was made to construct questions covering several areas, namely, road rules, good driving practices, car and driving safety, emergencies, defensive driving, equipment, driver improvement, and administrative practices. These were the same areas that were used as the basis for developing items on the written tests. However, the original areas had been delineated with the hope that it would be possible to use the written

Table 13. Numerical count of coefficient direction.

	A		B		C	
Question Number	-	+	-	+	-	+
1	0	1	4	4	0	1
2	1	2	1	0	2	2
3	0	3	0	0	3	2
4	0	1	0	2	2	1
5	1	1	0	2	2	0
6	0	0	0	0	4	3
7	1	2	4	0	1	0
8	0	1	4	1	0	0
9	2	0	0	6	0	2
10	5	1	0	3	0	1
11	1	0	2	4	5	1
12	2	6	0	1	2	0
13	1	7	4	0	4	0
14	3	2	0	1	1	2
15	3	0	2	0	1	3
16	0	1	0	3	1	3
17	0	0	2	0	1	5
18	0	0	1	0	2	0
19	1	1	0	0	2	1
20	2	3	0	1	1	1
21	3	2	2	0	1	0
22	1	0	0	4	2	0
23	1	3	2	2	2	1
24	3	2	2	4	0	1
25	0	0	3	3	2	1

Table 14. Question description and list of related driving record variables by time period, test form A.

1. Flashers: Stop (3)
2. Turn Signal: Acc Not Fault (1)
   Stop (2)
   Moving (3)
3. Direction When Turning: Speeding (1)
   Stop (1)
   Violations (1)
4. Backing: Moving (4)
5. School Bus: Acc Not Fault (1)
   Stop (2)
6. Slippery Stop
7. Drunk Driving: Acc (1)
   Stop (3)
   Reck (4)
8. Longer Stop: Speed (1)
9. Something in Road: Acc at Fault (3)
   Acc Not Fault (4)
10. Sleepy: Reck (1)
    Viol (1)
    Acc (1)
    Acc Not Fault (1)
    Stop (2,3)
11. Walk at Night: Stop (3)
12. Safe Turn: Stop (1)
    Reck (1,3)
    Acc at Fault (2,3)
    Acc Not Fault (2)
    Viol (3)
    Acc (3)
13. Skidding: Speed (2,3,4)
    Stop (2,3,4)
    Moving (4)
    Acc at Fault (4)
14. Lose License: Speed (3,4)
    Stop (4)
    Acc Fault (1)
    Acc (1)
15. Snowstorm Lights: Stop (1,2)
    Moving (2)
16. Driving Slow: Moving (3)
17. Passing
18. Breakdown
19. Bright Lights: Viol (2)
    Stop (3)
20. Interstate Accs: Speed (1)
    Viol (1)
    Reck (2,4)
    Acc Fault (2)
21. Stop: Acc (1,2)
    Stop (2)
    Viol (2,4)
22. Night Breakdown: Reck (2)
23. Bad Weather Speed: Viol (1,2)
    Moving (2)
    Acc Fault (3)
24. Off Road: Speed (1)
    Reck (3)
    Viol (3)
    Acc (2)
    Acc Not Fault (2)
25. Blind Intersection

Table 15. Question description and list of related driving record variables by time period, test form B.

1. Hill, Curve: Reck (2,3)
   Acc Fault (2,3)
   Accs (2,3)
   Acc Not Fault (2,3)
2. Street Parking: Moving (2)
3. Safe Driver
4. Safe Turn: Acc (2)
   Acc Not Fault (2)
5. Passing: Reck (2)
   Acc Fault (3)
6. Acc Behavior
7. Wet Brakes: Stop (2)
   Reck (4)
   Acc (4)
   Acc Not Fault (4)
8. Safe Driving: Stop (1)
   Moving (2,3)
   Reck (2,3)
9. Slick Roads: Moving (3)
   Reck (4)
   Acc Fault (4)
   Acc (1,4)
   Acc Not Fault (1)
10. Safe Passing: Speed (3)
    Reck (2)
    Viol (2)
11. 7 Pt. Record: Speed (3)
    Stop (2)
    Moving (2)
    Acc (3)
    Acc Not Fault (2,3)
12. Required Lights
13. Stuck in Sand: Reck (1)
    Acc Fault (4)
    Acc (4)
    Acc Not Fault (4)
14. Bicycle Law: Acc Fault (2)
15. Headlight Failure: Acc Fault (2)
    Acc (2)
16. Right-of-way: Stop (3)
    Acc Fault (4)
    Acc (4)
17. Brake Failure: Acc Fault (4)
    Acc (4)
18. Sleepy: Reck (3)
19. Interstate Car Trouble
20. Fog: Stop (1)
21. Lights for City/Night: Speed (2)
    Moving (4)
22. Interstate Acc Speeds: Stop (2)
    Acc Fault (4)
    Acc (4)
    Acc Not Fault (4)
23. Wrong Entry Indicators: Reck (3,4)
    Viol (2,3)
24. Skids: Speed (2,3)
    Stop (2,3)
    Reck (2)
    Viol (2)
25. Traffic Hazard: Stop (1)
    Moving (3)
    Reck (2)
    Acc Fault (3)
    Viol (1)
    Acc Not Fault (4)

Table 16. Question description and list of related driving record variables by time period, test form C.

1. Hauling: Stop (2)
2. Meeting Bright Lights: Moving (3)
   Acc Fault (1,3)
   Acc (1)
3. Blowout: Stop (2)
   Moving (2)
   Acc Fault (4)
   Viol (2)
   Acc (2)
   Acc Not Fault (3)
4. Speeding Up: Speed (1)
   Stop (2)
   Acc (3)
5. Signs: Reck (3)
   Acc Not Fault (1)
6. Engine Flooding: Speed (1)
   Stop (3,4)
   Acc Fault (3)
   Viol (1)
   Acc Not Fault (2,4)
7. Lose License: Stop (4)
8. Speeding Down Hill
9. Snow Starting: Speed (1)
   Reck (2)
10. Highway Acc: Acc Not Fault (4)
11. Leave Road: Stop (1)
   Moving (3)
   Reck (2,3)
   Acc Fault (4)
   Acc (4)
12. Changing Flat: Acc (3)
   Acc Not Fault (3)
13. Pass on Right: Stop (1) Acc (2)
   Moving (2)
   Acc at Fault (2)
14. Siren: Moving (2,3)
   Acc (2)
15. Long Trip Prep: Speed (3)
   Reck (2,3)
   Acc (2)
16. Pass: Reck (2,3,4)
   Acc (3)
17. Breakdown: Speed (3)
   Stop (2)
   Reck (3)
   Viol (2,3)
   Acc Not Fault (3)
18. Passing: Reck (3,4)
19. Rain: Speed (4)
   Acc (2)
   Acc Not Fault (2)
20. Right-of-way: Reck (1)
   Acc (2)
21. Head Lights: Stop (1)
22. Lose Lights: Stop (2)
   Moving (3,4)
   Acc Fault (2)
   Viol (2)
   Acc (2)
   Acc Not Fault (2)
23. Changing Lane: Moving (1,3)
   Acc Fault (4)
24. Skid: Stop (4)
25. Icy Stop: Speed (2)
   Moving (4)
   Reck (4)

Table 17. Question type distributed by test form.

	Total No.	A	B	C
Failures	11	18,22	7,15,17,19	3,6,12,17,22
Bad weather	9	6,15,23	9,13,20	9,19,25
Hazard	3	9,25	25	
Leave road	2	24		11
Skid	3	13	24	24
Situations	28	9	9	10
Passing	6	17	5,10	13,16,18
Turning	4	2,3,12	4	
Stop	2	8,21		
Speeding	3	16		4,8
Interstate	3	20	22	10
Right-of-way	2		16	20
Lights	5	19	12,21	2,21
Maneuvers	5	4	1,2	1,23
Procedures	30	10	9	11
Administration	10	1,5,7,14	6,11,14	5,7,14
Safe driving	5	11	3,8,23	15
Sleepy	2	10	18	
Practices & Rules	17	6	7	4

tests in a diagnostic sense. Such an approach was not indicated when the test analyses were completed, in part because the length of the test was not sufficient for such diagnostic purposes. The oral exam will be composed of even fewer items than the written tests, so that the question of diagnosis becomes even less feasible. Hence, for purposes of this analysis, content areas were first identified to equate questions over test forms. The questions were further grouped into only three major categories, roughly defined as follows: (1) situational questions which deal with predicaments in which a driver may find himself; (2) procedural questions which require the applicant to describe safe driving procedures involving turning, passing, etc.; and (3) administrative rules and general rules of good driving.

An item's predictive power is a measure of the correlation between a correct answer on that item and the dependent variable, i.e., violations and/or crashes on the driver record. The different types of questions seem to have differential predictive power. Table 18 shows the number of times that each type of question was used in the prediction of some driving record variable. Note that even though there are approximately an equal number of situational and procedural questions, the situational questions are used much more frequently in the prediction of previous driving records. This relationship remains the same in looking at just the prediction of accidents except that the numerical predictive power of the rules questions decreases with respect to the other question types. Some more specific types of questions appear to be more predictive of violations than crashes, such as questions concerning skidding, speed, lights, and safe driving practices.

Selection of Items for Recommended Test Forms.

One of the purposes of examining the data was to develop a more valid pictorial test. There are at least five different criteria which can be employed in selecting the better questions: difficulty level, item-total correlation, number of times the question was used in predicting previous driving record, item's predictive power regarding accidents, and the signs of the regression coefficients. One method that might be employed in order to combine these different criteria into a more useful decision-making instrument is non-metric multidimensional scaling (Shepard, Romney, and Nerlove, 1972). In this analysis each of the questions is ranked according to the five different criteria. The rank orderings are then analyzed using an internal unfolding analysis in two dimensions. Figures 2, 3, and 4 represent the spatial representation of this analysis for each test. The locations of the questions in the space are marked by the appropriate number.

Table 18. Question type predictive power for total driving record variables and accident variables.

Question Category	Question Type	Row	Standardized	Row	Standardized
Situational	Failures	33	3.0	17	1.55
	Bad Weather	23	2.56	9	1.0
	Hazard	8	2.67	4	1.3
	Leave Road	10	5.0	4	2.0
	Skid	9	3.0	1	.33
		83	2.96	35	1.5
Procedures	Passing	12	2.0	4	.6
	Turning	14	3.5	6	1.5
	Stop	6	2.0	2	.6
	Speed	4	1.33	1	.3
	Interstate	9	3.0	5	1.6
	Lights	8	1.6	2	.4
	Maneuvers	9	1.8	4	.8
		67	2.16	27	.87
Rules	Administrative	20	2.11	9	.9
	Safe Driving	9	1.8	1	.2
	Sleepy	6	2.0	2	1.0
		35	2.12	13	.75

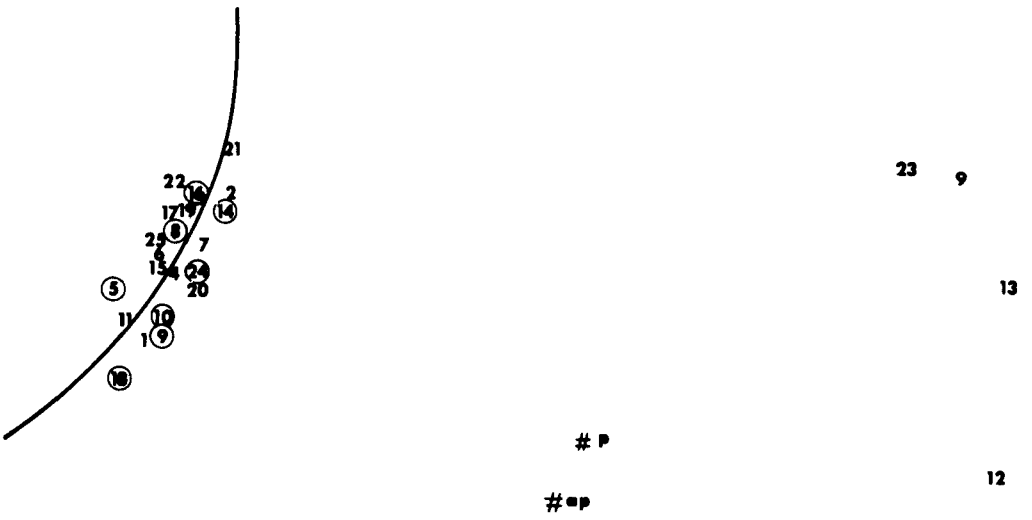

Figure 2. Question space for Test Form A.

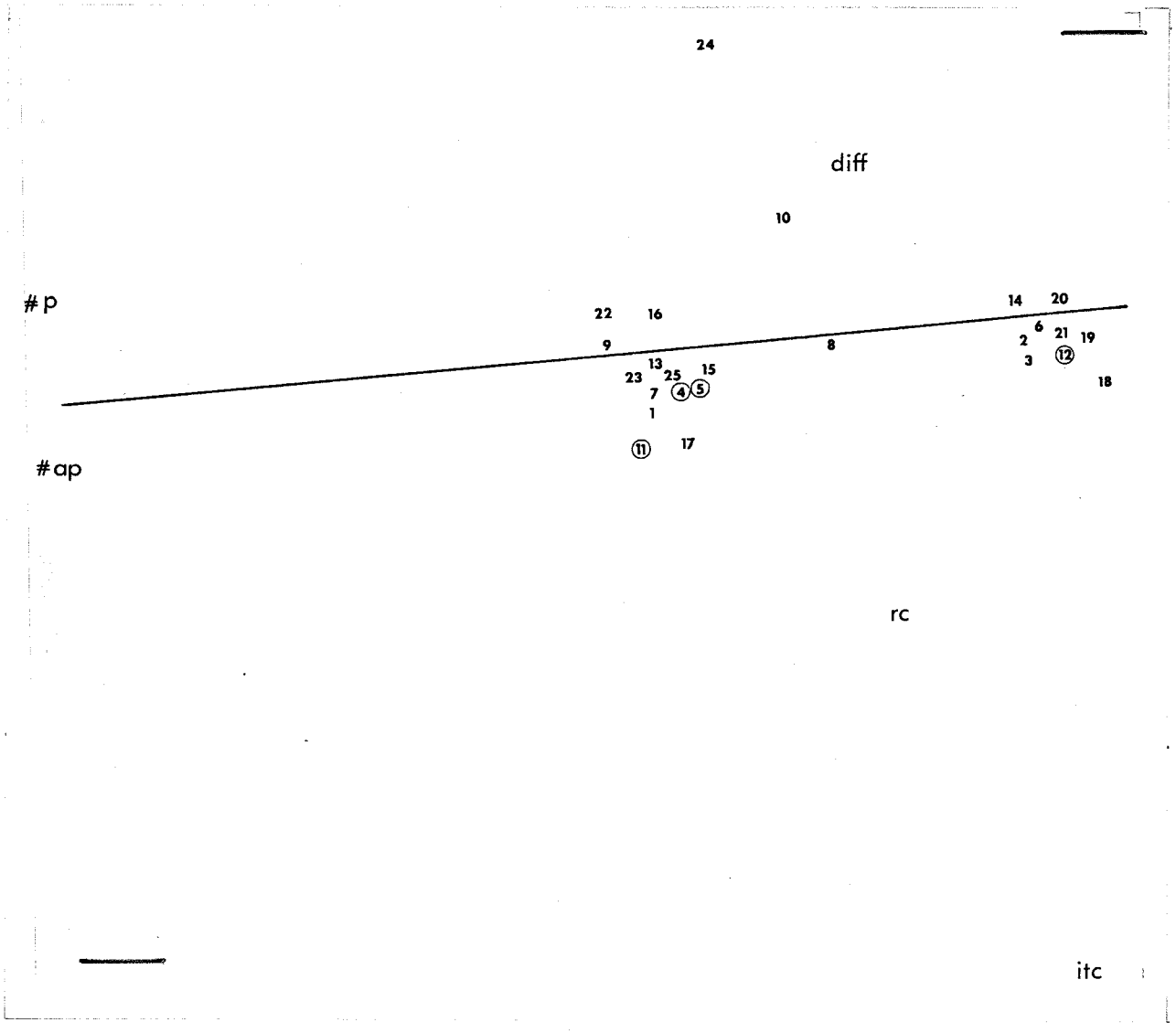

Figure 3. Question space for Test Form B.

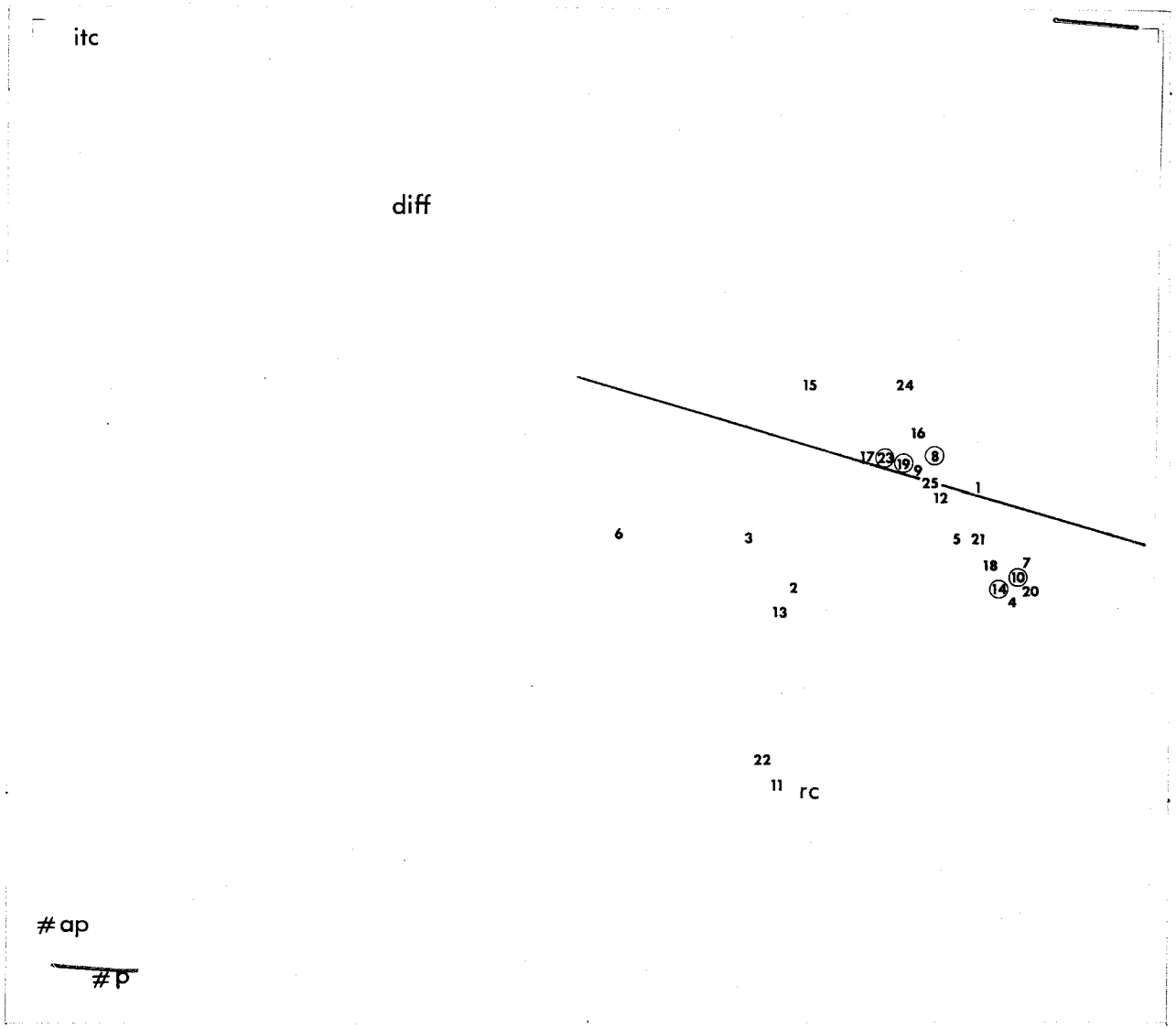

Figure 4. Question space for Test Form C.

The five criteria are also placed in the space. The closer a question number is to these points, the higher the question is ranked on those criteria.

This analysis provides some insights. For example, in test A there are clearly four questions (3, 12, 13, 23) which are detrimental to the examination for they clearly predict in an incorrect manner. The remaining questions seem to differ primarily in the difficulty levels and their item-total correlations. Test B separates the items into good predictors and poor predictors. There is also a good bit of variance in terms of the signs of the regression coefficients. Test C appears to cluster the questions into 3 categories: (1) good and appropriately directed predictors, (2) poor predictors with negative coefficients, and (3) poor predictors inappropriately signed.

However, the actual selection of questions to be retained should not be made on these results alone. This is because the five criteria are not equally important. The primary goal is to develop a test which provides some appropriate prediction about the previous driving record. One would prefer to have questions that were of no predictive power rather than questions that predicted inappropriately. To this end, lines have been drawn in an ad hoc manner which divide the questions into correctly and incorrectly directed prediction. Some caution must be taken, however, because the other criteria may have located some questions on the wrong side of the line. These questions are circled. These lines are useful in that only 28 percent, 16 percent, and 20 percent of the questions on the respective test forms seemed to be incorrectly located.

Table 19 contains a listing of the questions that might be employed in two shorter test forms. These are the questions that were either not predictive or predictive in an appropriate manner. These two tests have been constructed in an attempt to provide two parallel instruments. The questions were paired off according to subject matter as indicated in the table and then assigned to the two different test forms so that the expected test means would be approximately the same. The predicted test means of the two versions are approximately at the cutoff score for passing (.714 and .706). This control provides that neither of the new tests has more items which are too easy or too hard than the other test form. An attempt was also made to control for the mean item-total correlation.

Some further suggestions can be made. For example, questions B2, B6, and C20 from test Form 1 and questions C11, C18, and C22 from test Form 2 are too easy to be really good items. Two actions can be taken with respect to these questions. First, given this population, it

Table 19. Recommended composition of new test forms by type of question.

Type of Question	Test 1	Test 2
Situations	40	72
	53	62
	42	32
	56	22
	75	06
	69	38
	25	09
	50	61
Procedures	17	68
	63	21
	54	58
	46	15
	71	52
	70	26
	27	73
Rules	57	14
	10	43
	33	28
	55	48
	31	11

might be appropriate to leave these items in the tests and place them in the beginning. They would provide the applicant with a number of easy questions with which to become comfortable in the testing situation. These questions would also serve to give the applicant confidence. Second, one could revise the questions to create more reasonable alternatives and to raise the difficulty levels of the questions.

Finally, some concern must be stated for question A18. Although it is harmless in the sense that it makes no predictions, it is the only item in which the correct alternative did not dominate. It is suggested that the question be either revised or dropped from the test form. If A18 is dropped, then its similar item on Test Form 1, B19, can also be eliminated without loss of predictive power. This would provide for two parallel test forms with twenty questions each.

Table 20 lists two sets of items in the recommended order for two versions of the test, Alpha and Beta. The order has placed the easier items in appropriate locations but has otherwise equated the order of the two tests in terms of content covered. In addition, as indicated earlier, the predicted overall difficulty is approximately the same for the two test forms. These two versions of the test should provide two parallel driver license examinations that have meaningful relationships to driver records for applicants who do not read well enough to take the written test.

Ideally these two new versions of the pictorial oral exam should be field tested to insure that the test means and test reliabilities are acceptable and that the relationships between performance on individual items and driver records is in the appropriate direction.

SUMMARY

A need was identified for developing new driver license knowledge tests for use with applicants unable to read well enough to take the written version of the test. The new pictorial oral exams were based to a large extent on the efforts that had previously gone into the development of new written tests for North Carolina. However, each item was presented pictorially with the driver license examiner explaining the question and the answer choices to the applicant. There was considerable input from driver license examiners as well as others in the development of the pictorial items.

The new tests were introduced statewide in December of 1974, and special data were collected from the field in January of 1975.

Table 20. Recommended composition of new pictorial oral driver license examinations by item numbers.

Test Alpha	Test Beta
31	68
17	11
70	72
40	26
53	62
63	21
42	32
56	22
57	14
54	58
69	38
46	15
10	43
71	52
75	06
33	28
25	09
55	48
50	73
27	61
Predicted Mean Difficulty Level .715	.716
Mean Predicted Item Total Correlation .407	.404

Information on renewal applicants was linked to their prior driver records, and a variety of psychometric analyses were performed.

The results of the analyses included the following:

1. Applicants taking the oral version of the test differ markedly from those taking the written form. The oral exam applicants are, on the whole, older, much more likely to be male, much less likely to be white, and not as well educated as written exam applicants.

2. The mean scores on the three versions of the test were fairly close, averaging about 17 correct items. Nevertheless, a passing score required at least 18 correct items. It should be noted, however, that the mean scores on the new written tests were about 16 items correct, indicating that the new pictorial oral tests are slightly less difficult than the written rules examinations.

3. The test reliabilities are all at an acceptable level and are very close to those reported for the new written tests.

4. On the whole the item means indicate that the test items are at an acceptable level of difficulty.

5. The correlations between items and total test performance were generally acceptably high, while the inter-item correlations were generally low, indicating that on the whole the tests are assessing one general set of knowledge but from a variety of independent aspects.

6. The relationship between test performance and driver record variables appear stronger for reckless violations and for accidents than for other violation types or total violations.

7. Items were classified according to three categories, namely, situational questions dealing with predicaments in which a driver may find himself, procedural questions that require the applicant to describe safe driving procedures, and administrative rules and general rules of good driving. It was found that the situational

questions were more predictive of the driver record than were the other two types of questions.

8. Questions were finally considered on the basis of five criteria: difficulty level, item-total test correlation, number of times a question was used in predicting previous driving record, item's predictive power regarding accidents, and the direction of the relationship between item performance and driver record variables. The results of these evaluations led to the final selection of items.

9. Two final versions of the test are recommended, Form Alpha and Form Beta. These two tests have been equated for predicted overall difficulty level and for predicted item-total test performance. In addition, they each cover the same number of situational, procedural, and administrative questions. The order of items is the same for the different types of questions except for minor variations that were allowed in order to place unusually easy items at strategic points. Thus, for each version of the test the first, third, and final items are particularly easy. The two recommended versions of the test should provide two parallel driver license examinations that have meaningful relationships to driver records of applicants who do not read well enough to take the written tests. Ideally, these two new test forms should be field tested to insure that they indeed perform as well as would be anticipated on the basis of the present analyses.

REFERENCES

Creech, F.R., & Grandy, J. An analysis of the relationships between the renewal rules tests and the accident and violation histories of drivers in North Carolina. Princeton, NJ: Educational Testing Service, March 1974.

Magnusson, D. Test theory. Reading, MA: Addison Wesley, 1968.

McMichael, J., & Waller, P.F. The driver license rules test: The incidence and failure rates of the oral version compared to the written version. Chapel Hill: University of North Carolina Highway Safety Research Center, July 1973.

Shepard, R.N., Romney, A.K., & Nerlove, S.B. Multidimensional scaling. Vol. 1. New York: Seminar Press, 1972.

Stewart, J.R. An analysis of annual mileage self reported by renewal applicants. Chapel Hill: University of North Carolina Highway Safety Research Center, August 1975.

Tatsuoka, M. Multivariate analysis. New York: John Wiley, 1971.

APPENDIX A

Directions for Administering Pictorial Oral Examination

The applicant should preferably be seated opposite the examiner. The next most desirable arrangement would be for the applicant to be seated catercorner (diagonally) to the examiner. With the applicant seated opposite the examiner, the written material for each question will be in the appropriate place for the examiner to see it.

The examiner should introduce the examination in the following way: "I am going to ask you some questions, and for each question there will be several pictures or a single picture with several parts. I will explain the question to you, and then you are to point to the picture or the part of a picture that shows the right answer. Do you have any questions?"

If the applicant has no questions, the examiner says, "Let's try one." He then proceeds with the sample question as explained below.

If the applicant has questions, the examiner should handle them and then say, "Let's try one of these questions now."

The examiner will turn to the sample in the beginning of the examination notebook. He will state clearly the first part of the question and then point to the appropriate picture as he gives each answer choice. He should then pause for the applicant's response.

IF THE APPLICANT THEN POINTS TO THE CORRECT PICTURE, the examiner should tell him, "Good. Now let's try some more." The examiner will then proceed to one of the test forms.

IF THE APPLICANT POINTS TO THE WRONG ANSWER, the examiner should say, "The right answer is this one. Highway accidents happen most often at intersections. Now let's try some more questions." The examiner will then proceed to one of the test forms.

IF THE APPLICANT DOES NOT ANSWER, the examiner should repeat the question, pause momentarily, and then point to the correct answer himself, saying, "This picture shows the right answer. Highway accidents happen most often at intersections. Let's try this question one more time." He should repeat the sample question one more time, giving the applicant a chance to answer. If the applicant still does not answer or does not give the right answer, the examiner should say, "Let's try some more questions." He should then proceed to one of the three test forms. In no case should the sample question be explained more than twice.

For each question thereafter the examiner should read the first part of the question clearly and point to each picture response as he goes over the answer choices. For each picture he should point out the relevant parts as he gives the answer choice. As the test proceeds, the examiner will need to record the answers given by the applicant so that a final score can be calculated.

APPENDIX B

Correlation Matrices for Test Forms A, B, and C

PROGRAM NUMBER 2.
EXECUTING CORRELATIONS PROGRAM
ADDRESS OF MEMORY WORK AREA IS 06E800. LENGTH OF MEMORY WORK AREA IS 028800 (022272).

SAMPLE SIZE = 113 TEST FORM C

CORRELATION MATRIX

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1.00000													
2	0.05730	1.00000												
3	0.22191	0.14886	1.00000											
4	0.02565	0.12226	0.17329	1.00000										
5	0.12230	0.06380	0.02071	0.12596	1.00000									
6	0.17727	0.14861	0.30665	0.10351	-0.00288	1.00000								
7	0.30979	0.04368	0.10872	0.10307	-0.06553	0.10462	1.00000							
8	0.11820	0.08586	0.18675	0.28993	0.08090	0.37074	0.13100	1.00000						
9	0.03731	0.00783	0.03909	0.17212	0.05419	0.25916	0.08548	0.24270	1.00000					
10	-0.06604	0.23520	0.05557	0.06235	-0.13227	0.29650	0.16792	0.16033	0.11763	1.00000				
11	-0.00541	-0.11031	0.31829	-0.00965	-0.00352	0.21344	0.08640	0.13238	0.16007	0.05017	1.00000			
12	0.15662	0.13085	0.15136	0.21147	-0.03964	0.21812	0.12002	0.25798	0.23409	0.23274	0.22968	1.00000		
13	0.19432	0.24479	0.17221	0.08233	0.04623	0.06613	0.00630	0.05593	0.16592	0.04418	0.15430	0.19675	1.00000	
14	-0.01449	0.05629	0.21842	-0.04429	-0.10085	0.25867	-0.03008	0.14982	0.19323	0.18200	0.20125	0.14962	0.13654	1.00000
15	0.13745	0.22632	0.34277	0.03995	0.05384	0.18866	0.06335	0.17948	0.16071	0.21416	0.23834	0.36351	0.21188	0.29777
16	0.09804	0.06100	0.26581	0.08499	0.03840	0.21614	0.07248	0.17125	0.12492	0.00686	0.25555	0.21044	0.28595	0.19558
17	0.19683	0.14050	0.19022	0.17933	0.01137	0.23400	0.18664	0.21080	0.25365	0.08921	0.08292	0.21932	0.07042	0.01170
18	0.23864	0.13783	0.10424	0.01131	-0.13504	0.19019	0.05705	0.11352	0.01723	0.02334	-0.05140	0.14642	0.13032	-0.11062
19	0.28751	0.26342	0.29165	0.13465	0.22196	0.17943	-0.01219	0.22231	0.11154	0.06371	-0.02840	0.14284	0.11367	0.26771
20	0.27342	0.20467	0.02819	0.04571	-0.05463	0.08883	0.03899	0.03396	-0.15450	0.05273	0.02240	0.06066	0.20561	0.06000
21	0.05340	0.01834	0.12908	0.11592	0.09215	0.03400	-0.14939	-0.06747	0.14403	-0.12278	0.16215	0.07890	0.01235	0.01365
22	0.07716	0.13783	0.23022	0.06690	-0.13504	0.19019	0.16450	0.11352	-0.04360	0.18481	0.11696	0.14642	0.01660	0.08780
23	0.23274	0.16896	0.19590	0.13287	0.07721	0.21812	0.08203	0.16773	0.14807	0.04244	0.11062	0.02165	0.11634	0.21977
24	0.07530	0.30833	0.19619	0.14693	-0.02596	0.19859	0.13845	0.20911	0.26112	0.16487	0.17294	0.23941	0.21770	0.10300
25	0.05340	0.15064	0.18063	0.16140	0.00199	0.08326	0.20230	0.03697	0.14403	0.09744	0.16215	0.12607	0.24497	0.01365

	15	16	17	18	19	20	21	22	23	24	25
15	1.00000										
16	0.31960	1.00000									
17	0.23745	0.14174	1.00000								
18	0.03856	0.09655	0.16591	1.00000							
19	0.23834	0.17956	0.12930	0.06600	1.00000						
20	0.06605	0.16036	0.15051	0.23765	0.19954	1.00000					
21	-0.00589	0.15401	0.05470	0.01004	0.23403	0.05245	1.00000				
22	0.21283	0.16561	-0.05145	0.10317	0.17900	0.12622	-0.05668	1.00000			
23	0.11705	0.25927	0.21932	0.26172	0.22274	0.13945	0.12607	0.08877	1.00000		
24	0.29899	0.28220	0.21046	0.15428	0.25419	0.15179	0.14000	0.28996	0.23941	1.00000	
25	0.04165	0.15401	0.27702	0.14346	0.14156	-0.03873	0.12657	0.27689	0.36194	0.25102	1.00000

EXECUTE TIME FOR CORRELATIONS PROGRAM IS 0.64 SECONDS.
* *

PROGRAM NUMBER 2.
EXECUTING CORRELATIONS PROGRAM
ADDRESS OF MEMORY WORK AREA IS 094000. LENGTH OF MEMORY WORK AREA IS 028800 (022272).

SAMPLE SIZE = 145 TEST FORM B

CORRELATION MATRIX

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1.00000													
2	0.12809	1.00000												
3	0.11477	0.21142	1.00000											
4	0.10322	0.07059	0.18732	1.00000										
5	0.13095	0.27112	0.36102	0.16290	1.00000									
6	0.12809	0.23885	0.30737	0.15450	0.21575	1.00000								
7	0.21630	0.11392	0.09483	0.16157	0.21658	0.07135	1.00000							
8	0.20272	0.16223	0.26479	0.28801	0.26087	0.07685	0.15474	1.00000						
9	0.24129	0.04963	0.12287	0.17137	0.21032	0.13331	0.19968	0.24219	1.00000					
10	0.11153	0.08872	0.22966	0.15528	0.32154	0.22805	0.16576	0.33135	0.05895	1.00000				
11	0.21610	0.20828	0.10612	0.26667	0.29542	0.12404	0.19908	0.24381	0.25944	0.24970	1.00000			
12	0.17805	0.30737	0.20170	0.15559	0.31915	0.16345	0.22364	0.13562	0.15452	0.19453	0.32913	1.00000		
13	0.00322	0.20568	0.05631	0.14374	0.21499	0.15116	0.27157	0.16881	0.08160	0.14564	0.23446	0.09755	1.00000	
14	-0.02660	0.18880	0.16965	0.17076	0.20847	0.09520	-0.04007	0.09449	0.04364	0.09216	0.17573	0.13425	0.03329	1.00000
15	0.22029	0.18880	0.13425	0.04696	0.20847	0.09520	0.17980	0.22048	0.13624	0.05789	0.17573	0.24044	0.07351	0.17143
16	0.07981	0.01568	0.14932	-0.00638	0.11231	0.19058	0.08015	0.16481	0.16257	0.16133	0.10493	0.06113	0.20321	0.12607
17	0.16054	0.03980	0.20170	0.07166	0.18511	0.20828	0.11426	0.21546	0.14831	0.12634	0.18885	0.26541	0.05343	0.20681
18	0.00243	0.20118	0.19778	0.07094	0.18888	0.26794	0.15636	0.07189	0.13938	0.16989	0.09141	0.24827	0.23703	0.17661
19	0.07730	0.05744	0.12241	0.18277	-0.00678	0.22627	0.12218	0.12879	0.19482	0.13243	0.13428	0.16498	0.08207	0.14030
20	0.12456	0.16095	0.19862	0.14048	0.29442	0.06996	0.27777	0.07349	0.04367	0.21847	0.13355	0.19862	0.12163	0.12962
21	0.02244	0.20165	0.25358	0.12958	0.22719	0.14222	0.09794	0.10400	0.14572	0.08973	0.05008	0.20863	0.11415	0.22830
22	0.26922	0.07059	0.21905	0.05664	0.19952	0.28035	0.16157	0.00565	0.11603	0.12456	0.15524	0.06040	0.07162	0.01601
23	0.00579	0.02343	0.05427	0.09201	0.09843	0.10836	0.04521	0.08575	0.01739	0.18764	0.09821	0.21487	-0.02769	0.10803
24	0.20378	0.04503	0.02308	-0.05330	0.08793	0.13095	0.06383	0.07518	0.15809	0.00521	0.13457	0.08806	-0.11230	0.01749
25	0.11944	0.05609	0.23627	0.23029	0.29040	0.27241	0.17503	0.24460	0.18716	0.18154	0.16762	0.13811	0.01180	0.06714

	15	16	17	18	19	20	21	22	23	24	25
15	1.00000										
16	0.04004	1.00000									
17	0.11358	0.22107	1.00000								
18	0.17661	0.21919	0.18006	1.00000							
19	0.01575	0.27718	0.17165	0.31696	1.00000						
20	0.12962	0.05537	0.25439	0.19844	0.15004	1.00000					
21	0.18445	0.09908	0.20794	0.48052	0.18582	0.32512	1.00000				
22	0.07791	0.07073	0.09952	0.11509	0.14555	0.08030	0.01166	1.00000			
23	0.23335	0.15209	0.09821	0.19572	0.22866	0.02416	0.17150	0.09201	1.00000		
24	0.11256	0.25678	0.16309	0.08231	0.18032	0.03102	-0.00444	0.11717	0.05256	1.00000	
25	0.09905	0.07185	0.19635	0.13847	0.07915	0.05007	0.12887	0.20167	0.25069	0.03354	1.00000

EXECUTE TIME FOR CORRELATIONS PROGRAM IS 0.75 SECONDS.
* *

SAMPLE SIZE = 232 TEST FORM A

CORRELATION MATRIX

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1.00000													
2	0.02753	1.00000												
3	0.13072	0.25459	1.00000											
4	0.09569	0.06918	0.10661	1.00000										
5	0.16097	0.19377	0.26845	0.14628	1.00000									
6	0.06689	0.16168	0.16647	0.18517	0.08345	1.00000								
7	0.17128	0.17568	0.16423	0.16040	0.18884	0.10025	1.00000							
8	0.08502	0.22628	0.24434	0.16457	0.14645	0.21984	0.03790	1.00000						
9	0.00316	0.11225	0.19108	0.04463	0.01671	0.12193	0.11750	0.09904	1.00000					
10	0.10830	0.07705	0.11402	0.15899	0.10257	0.13895	0.23002	0.12581	0.20164	1.00000				
11	0.05708	0.21562	0.06334	0.04506	0.07712	0.11177	0.14957	0.15145	0.06124	0.14364	1.00000			
12	0.15748	0.20249	0.11098	-0.01882	0.10235	0.01197	0.11510	0.19207	-0.01944	0.10568	0.18722	1.00000		
13	-0.00246	0.08589	0.07512	0.08427	0.16378	0.05384	0.14840	0.02721	0.02261	0.07304	0.10064	-0.05384	1.00000	
14	0.07481	0.22031	0.13072	0.10025	0.22421	0.10139	0.25302	0.14434	-0.02248	0.14783	0.14783	0.17908	0.13536	1.00000
15	0.05453	0.13042	0.10820	0.11622	0.11072	0.01634	0.10366	0.17090	0.01344	0.10714	0.10714	0.06104	0.08366	0.18246
16	0.12851	0.20739	0.23298	0.19125	0.16013	0.10358	0.19509	0.19412	0.02720	0.18179	0.13106	0.17412	0.10355	0.27274
17	-0.02996	0.26770	0.19460	0.20682	0.16983	0.07719	0.21819	0.07811	0.11775	0.18599	0.15421	0.06199	0.19127	0.27040
18	0.03414	0.17444	0.03627	0.05554	-0.00377	0.04898	0.01721	0.03292	0.03535	0.01838	-0.00827	0.04831	0.08175	0.14921
19	0.01497	0.31269	0.21900	0.16611	0.20033	0.16224	0.19246	0.19895	0.17641	0.22998	0.22998	0.08506	0.15651	0.19477
20	0.09451	0.08965	0.13193	0.18301	0.08675	0.08801	0.04221	0.20093	-0.01059	0.16613	0.14028	0.10070	0.02032	0.13667
21	0.01080	0.18191	0.24045	0.08120	0.19493	0.08724	0.10025	0.20111	0.12193	0.24767	0.11177	0.09134	0.19488	0.13879
22	0.13838	0.27945	0.22929	0.10907	0.28543	0.08093	0.27929	0.23444	0.11062	0.14548	0.14548	0.12405	0.29319	0.15136
23	0.00396	0.21559	0.22740	0.06692	0.17723	0.12006	0.16448	0.16813	0.18802	0.21422	0.15347	0.10171	0.14210	0.14232
24	0.00089	0.24410	0.12240	0.03064	0.02950	0.14829	0.14779	0.06587	0.19528	0.11345	0.08347	0.09244	0.19348	0.12285
25	0.07059	0.22933	0.11991	0.06814	0.13943	0.08529	0.20470	0.11016	0.14976	0.22386	0.19182	0.10182	0.12878	0.21594

	15	16	17	18	19	20	21	22	23	24	25
15	1.00000										
16	0.12215	1.00000									
17	0.12984	0.19459	1.00000								
18	0.10008	0.09578	0.04235	1.00000							
19	0.34276	0.33089	0.29682	0.08362	1.00000						
20	0.10801	0.23762	0.00970	0.18377	0.20306	1.00000					
21	0.05504	0.27020	0.19316	0.14626	0.26116	0.25785	1.00000				
22	0.03342	0.19933	0.09378	0.13585	0.20196	0.12587	0.18342	1.00000			
23	0.12153	0.25080	0.25120	-0.00150	0.36647	0.10218	0.23095	0.27966	1.00000		
24	0.12618	0.16669	0.16099	0.12062	0.24336	0.15915	0.21394	0.15007	0.14423	1.00000	
25	0.16179	0.22848	0.19351	0.10360	0.15844	0.15657	0.20224	0.19809	0.18073	0.29732	1.00000

EXECUTE TIME FOR CORRELATIONS PROGRAM IS 0.98 SECONDS.

* *

Original photograph

Original artwork