HIGHWAY RESEARCH BOARD

Bulletin 60

Road-User Characteristics

National Academy of Sciences—
National Research Council

publication 244

HIGHWAY RESEARCH BOARD

Bulletin 60

Road-User Characteristics

Presented at the
THIRTY-FIRST ANNUAL MEETING
January 1952

1952
Washington, D. C.

TRAFFIC AND OPERATIONS DEPARTMENT

Wilbur S. Smith, Chairman
Associate Director, Bureau of Highway Traffic
Yale University

COMMITTEE ON ROAD USER CHARACTERISTICS

Forbes, Theodore W., Chairman; Executive Secretary, Committee on Highway Safety Research, 2101 Constitution Avenue, Washington 25, D. C.

Allgaier, Earl, Research Engineer, Traffic Engineering and Safety Department, American Automobile Association, Mills Building, Washington 6, D. C.

Baker, J. Stannard, Director of Research, Traffic Institute, 1704 Judson Avenue, Evanston, Illinois

Case, Howard W., Department of Engineering, University of California, Los Angeles 24, California

Crandall, F. Bruce, Traffic Engineer, Oregon State Highway Commission, State Highway Building, Salem, Oregon

Deal, George V., Institute of Public Safety, Pennsylvania State College, State College, Pennsylvania

Eliot, William G. 3d, Highway Engineer, Bureau of Public Roads, U. S. Department of Commerce, Washington 25, D. C.

Flanakin, H. A. Mike, Acting Director, Engineering Experiment Station, College of Engineering, Louisiana State University, Baton Rouge 3, Louisiana

Gordon, Nathan, Highway Engineer, Bureau of Public Roads, U. S. Department of Commerce, Washington 25, D. C.

Gravelle, Gordon K., Transportation Engineer, Automotive Safety Foundation, 700 Hill Building, Washington 6, D. C.

Holmes, Edward H., Chief, Highway Transport Research Branch, Bureau of Public Roads, U. S. Department of Commerce, Washington 25, D. C.

Kraft, Merwyn A., Director, Department of Personnel and Accident Prevention, American Transit Association, 292 Madison Avenue, New York 17, N. Y.

Lauer, A. R., Professor of Psychology, Director, Driving Research Laboratory, Iowa State College, Station A, Ames, Iowa

McFarland, Ross A., Harvard School of Public Health, 695 Huntington Avenue, Boston 15, Massachusetts

Matson, Theodore M., Director, Bureau of Highway Traffic, 311 Strathcona Hall, Yale University, New Haven 11, Connecticut

Moffie, Professor D. J., Head, Department of Psychology, State College of Agriculture & Engineering, University of North Carolina, Raleigh, North Carolina

Moseley, Alfred L., Highway Transport Research, Harvard School of Public Health, 695 Huntington Avenue, Boston 15, Massachusetts

Neal, Harry E., Chief Engineer, Division of Traffic & Safety, Ohio Department of Highways, Ohio Departments Building, Columbus 15, Ohio

Uhlaner, Julius E., Chief, Recruitment Section and Classification Research Unit, Personnel Research Section, AGO, Department of the Army, Pentagon Building, Washington 25, D. C.

Volk, Wayne N., Traffic Engineer, State Highway Commission of Wisconsin, 1 West Wilson Street, Madison 2, Wisconsin

Wilson, Clark L., Management & Marketing Research Corporation, 721 West Jefferson Boulevard, Los Angeles 7, California

Wynn, F. Houston, Bureau of Highway Traffic, 311 Strathcona Hall, Yale University, New Haven 11, Connecticut

Yost, Colonel Light B., Director, Field Operations Section, Distribution Staff, General Motors Corporation, General Motors Building, 3044 West Grand Boulevard, Detroit 2, Michigan

Young, J. C., Traffic Engineer, California Division of Highways, P. O. Box 1499, Sacramento 7, California

CONTENTS

Road-User Characteristics

DEVELOPMENT of CRITERIA of SAFE MOTOR-VEHICLE OPERATION

J. E. Uhlaner, Chief, Research Operations II
Leon G. Goldstein, Research Associate
N. J. Van Steenberg, Statistical Advisor
Personnel Research Section, AGO
Department of the Army

THE continuing cost of motor-vehicle accidents in the U. S. Army led to the establishment in 1949 of a broad program of research to investigate psychological factors involved in safe motor-vehicle operation by the military. The consensus of investigators of motor-vehicle accidents is that only a small proportion of accident variance can be ascribed to faulty equipment of the vehicles involved or to hazards inherent in the roadway. The vast majority of accidents are due to the human operators. So while no efforts should be spared to perfect the safety features of roads and of the cars, reduction in accident rate in the military is probably best effected by the proper selection, training, and managing of military drivers. Therefore, the emphasis in this research program is placed on the psychological factors involved in safe and effective vehicle operation.

Within the over-all research program, a number of projects were set up; some have been completed, some are now nearing completion, and some are still to be initiated. These projects were designed to examine various psychological aspects of motor-vehicle operation in order to promote safety and effectiveness in army motor transportation.

Examples of the projects being carried out under this research program will serve to indicate the scope of the research effort. One project is concerned with "Personnel Management Factors in Vehicle Safety" and was executed under contract with Richardson, Bellows, Henry, and Company, Inc., Harold A. Edgerton acting as the principal investigator. The purpose of that research project was to study the relationship of administrative practices to safety performance of army motor-vehicle pools. Two other projects are concerned with the development of experimental predictors of safe and effective vehicle operation. Some of these predictors are being developed by the Personnel Research Section of The Adjutant General's Office, while others are being developed and validated under contract by Iowa State College with A. R. Lauer as principal investigator. The details of the above projects will not be reported in this paper. This report deals with the first study under this program which is concerned with the development of a criterion measure of safe and effective vehicle operation which will be more reliable and meaningful in the case of the military driver than a road test or accident rate by itself. An acceptable criterion was necessary in order to evaluate tests, procedures, and other techniques that would be investigated under this research program.

INADEQUACY OF ACCIDENT RECORDS

Although most of the civilian studies on safe driving in the civilian situation have used accident records as the criterion, only a very few have attempted to evaluate the stability of accident records. The most comprehensive data on the subject are furnished by the Connecticut study of the Bureau of Public Roads (8) which used the 6-year records (1931-36) of over 29,000 licensed drivers. In Table 1 the distribution of accidents in the second 3-year period is compared with that of the same drivers in the first 3-year period.

TABLE 1

ACCIDENTS OF GENERAL DRIVERS IN CONNECTICUT IN YEARS 1934-36 COMPARED WITH THOSE OF THE SAME DRIVERS IN YEARS 1931-33[a]
(from Bureau of Public Roads Study)

Accidents per operator, 1934-36	Accidents occurring to same operators, 1931-33: 0	%	1	%	2	%	3	%	4	%	Total number of operators	%
0	23881	(91)	2386	(83)	275	(77)	22	(71)	5	(50)	26569	(90)
1	2117	(8)	419	(15)	64	(18)	5	(16)	4	(40)	2609	(9)
2	242	(.9)	57	(2)	12	(3)	2	(6)	0	(0)	313	(1)
3	17	(.6)	9	(.3)	5	(1)	2	(6)	1	(10)	34	(.1)
4	2	(.0)	3	(.1)	1	(.3)	0	(0)	0	(0)	6	(.02)
Totals	26259	(89)	2874	(9.7)	357	(1.2)	31	(.1)	10	(.03)	29531	(100)

a/- Percentages inserted by present authors.

This table shows a definite tendency of those who were accident free in the first period to be accident free in the second period, and a progressive likelihood of those who had accidents in the first period to have accidents in the second period. However, the correlation (tetrachoric) of accident experience in the two periods comes to 0.24.

Farmer and Chambers (3) correlated accident experience in separate years for four groups of British public-transportation drivers. These correlations are shown in Table 2.

TABLE 2

PRODUCT-MOMENT CORRELATIONS BETWEEN ACCIDENTS IN SEPARATE YEARS OF EXPOSURE
(from Farmer and Chambers)

Correlation between accidents in years:	Group A 166 Bus Dr.	Group B 398 Bus Dr.	Group C 86 Bus and Trolley Dr.	Group D 67 Trolley Dr.
1 + 2	.298	.182	.235	.071
1 + 3	.235	---	.063	.058
1 + 4	.177	---	.281	.127
1 + 5	.274	---	---	---
2 + 3	.328	---	.078	.225
2 + 4	.176	---	.195	.251
2 + 5	.265	---	---	---
3 + 4	.212	---	.016	.296
3 + 5	.273	---	---	---
4 + 5	.224	---	---	---

For the 166 bus drivers in Group A, the correlations between accidents in the first year and accidents in increasing subsequent periods were:

Between 1st and 2nd year	r = 0.298
Between 1st and 2-3 years	r = 0.327
Between 1st and 2-4 years	r = 0.339
Between 1st and 2-5 years	r = 0.375

Brown and Ghiselli (2) estimated the reliability coefficients from correlations between the number of accidents on the odd and even months over a period of 18 months corrected by the Spearman-Brown formula. For 59 trolley-car motormen in California the estimates for different types of accidents are:

Collision with pedestrian	r = 0.46
Collision with trolley cars	r = 0.19
Collision with motor vehicles	r = 0.42
All collision accidents	r = 0.42

Bransford (1) correlated accident frequency during the year and a half after administration of driver tests against accident frequency during a variable period prior to testing. The correlation between accident rates, for a group of 481 drivers in Washington, D. C., was 0.184.

Slocombe (5) correlated yearly accident rates of 260 motormen of the Boston Elevated Railroad over a period of 4 years. The correlations between different yearly periods were as follows:

1st and 2nd year	r = 0.51
1st and 3rd year	r = 0.43
1st and 4th year	r = 0.38
2nd and 3rd year	r = 0.41
2nd and 4th year	r = 0.38
3rd and 4th year	r = 0.43

CONDITIONS PECULIAR TO THE ARMY SITUATION

It was hardly to be expected that accident records in the army situation would be more stable than those of these civilian studies. On the contrary, they might well be less stable, for the following reasons:

1. The average mileage of army drivers per year is approximately 12,000 to 15,000 mi. Even if a driver were kept on duty at the same pool during a full enlistment and his records were therefore available, this extent of exposure would hardly be sufficient to yield a reliable measure of safety of driving behavior. In addition, there may be vast observable differences in performance of two drivers after 15,000 mi. of driving but with both drivers showing zero accident rates.

2. The distribution of accidents is curtailed because of the policy of removing a driver from driving duty when he has had a second or third accident.

3. Driving conditions from pool to pool vary greatly in terms of vehicles used, mission, supervision, climate, terrain, density of traffic, night versus day driving, etc. Accident rates would thus be contaminated with uncontrolled variables.

INADEQUACY OF ROAD TESTS

Previous studies of the Personnel Research Section (6) furnish considerable data on the reliability of road tests. Four kinds of estimates of reliability were made: (1) the relationship between separate check-list items and general ratings on the road test (N = 1717) ranged between r = 0.22 and r = 0.57; (2) the split-half reliability of the Road Test Check list, using the Spearman-Brown formula, was computed as r = 0.82 (N = 155); (3) the correlation of scores on the Road Test Check List given by different examiners at different times (same 155 cases) was r = 0.53; and (4) the reliability of the general ratings on the road test was computed by correlating two series of ratings of 127 men made on the same day by several specially trained examiners. This yielded a coefficient of r = 0.72.

Although these reliabilities are generally higher than those of accident records, the road test suffers from one serious objection as a measure of safe driving behavior, regardless of its reliability. Driving behavior during such a test, under the surveillance of one whom the driver recognizes as an examiner, may be expected to be different from driving behavior under ordinary conditions.

SPECIFIC APPROACH OF THIS PROJECT

In the light of the shortcomings of both accident records and road tests as criteria for the evaluation of instruments for the selection of safe drivers, the decision was made to explore the possibility of assessing the driving behavior of army drivers on the basis of the observations and pooled judgments of their supervisors and associates. In view of the practical considerations of administration, and in view of the generally high intercorrelations obtained among rating scales and check lists, it was decided to develop a criterion instrument of about four simple rating scales and a driving habit check list of about 15 items.

DEVELOPMENT OF RATING SCALES

Eleven experimental scales were constructed from which the final four were to be selected. Eight aspects of driving behavior were postulated and a 15-point rating scale was designed to measure each. Two additional scales were designed to be used as psychological suppressors and another scale was designed for an over-all safe-unsafe rating. The lead questions of the 11 scales are given below.

1. How often does he have near accidents?

2. How well does he react to sudden changes of traffic conditions?

3. How much does "temper" or "nerves" affect his driving?

4. How well does he know his own limitations - like poor eyes, slowness, lack of skill, etc. - and drive according to what he knows he can do?

**5. How safe a driver is he?

6. What is his attitude toward safety when he drives?

7. How well does he keep his mind on his driving?

*8. How well do you like him?

9. How skillful is he in handling a vehicle?

*10. How does he rate on appearance and military bearing?

11. How well does he take care of his vehicle?

*Designed to be used as psychological suppressors.
**Designed to obtain an over-all safe-unsafe rating.

Each scale was divided into 5 sections (3 scale-points each) with a verbal definition of each section printed therein. On 10 of the 11 scales, cartoons designed to illustrate the intent of the lead question were printed at the "good" and "bad" ends of the continuum. The 11 scales were printed in booklet form, preceded by a practice rating scale. A sample of one of the scales is shown in Figure 1. The booklet was arranged so that the rater could not see, as he worked, what ratings he had given on the previous scale.

ADMINISTRATION OF EXPERIMENTAL SCALES

A trial run of this booklet was conducted during July 1950 in seven motor pools in the First Army Area. The motor pools visited were at Fort Jay, N. Y.; West Point, N. Y.; Manhattan, N. Y.; and Fort Dix, N. J. Rosters of not more than 20 drivers each were drawn up in such a manner that familiarity with one another's driving behavior was maximal. In rating sessions conducted with drivers and with their supervisors separately, a total of 200 drivers were each rated by 2 to 10 (mean of 4.8) supervisors, and by 5 to 26 (mean of 12.5) associate drivers. A total of 23 sessions were conducted. All ratings were anonymous.

ANALYSIS OF RESULTS

In order to select the four scales which would best measure safe driving behavior, the following possible bases of selection were investigated: (1) reliabilities of the 11 scales; (2) correlations of mean ratings on the scales with an index of accident responsibility; and (3) the results of a factor analysis of the ratings which was intended to identify those scales with high loadings on that orthogonal factor (or factors) which represents the variance of most of the scales and have low loadings on the orthogonal factor identified with the suppressor scales.

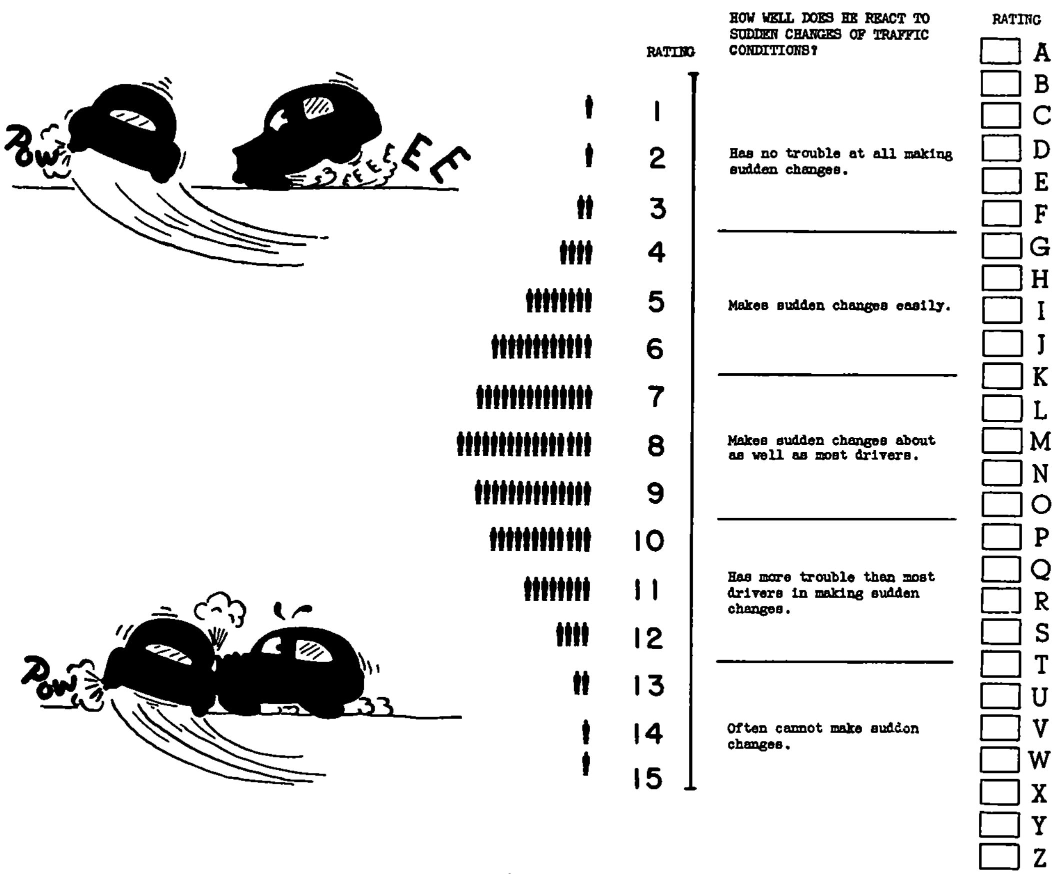

Figure 1.

RELIABILITY OF THE EXPERIMENTAL SCALES

The reliability of each scale was estimated by means of a modification of the Horst formula. (4):

$$\text{rel.} = 1 - \frac{\sum \frac{n_i \sigma_i^2}{n_i - 1}}{\sigma_{M_i}^2 \sum n_i}$$

where

- n_i is the number of ratings for driver *i*
- σ_i is the standard deviation of these ratings for driver *i*
- σ_{M_i} is the standard deviation of the means for the *N* drivers

The estimates of reliability and the means and standard deviations of mean ratings on each scale are shown in Table 3 for supervisors' and for associates' ratings.

TABLE 3.

MEANS AND STANDARD DEVIATIONS OF MEAN RATINGS ON 11 DRIVER RATING SCALES, AND THE RELIABILITIES OF THE 11 SCALES

	Supervisors' Ratings (181 Ratees)				Associates' Ratings (189 Ratees)		
Scale	M	S.D.	Rel.	Rel. corrected for no. of raters	M	S.D.	Rel.
1	4.78	2.21	.85	0.94	4.12	1.49	.94
2	5.56	1.93	.81	0.93	4.58	1.58	.94
3	5.47	1.82	.79	0.92	4.46	1.49	.94
4	5.84	2.04	.83	0.94	4.53	1.55	.93
5	5.87	1.98	.85	0.94	4.96	1.67	.94
6	5.92	1.92	.83	0.94	4.58	1.55	.93
7	6.01	1.93	.84	0.94	4.95	1.44	.93
8	6.00	1.45	.80	0.92	4.70	1.29	.93
9	5.55	1.70	.83	0.94	4.88	1.52	.93
10	5.68	2.03	.85	0.94	4.64	1.44	.93
11	5.65	1.92	.83	0.93	4.64	1.48	.93

In order to ascertain whether the lower reliabilities of the supervisors' ratings were attributable to the smaller number of supervisory ratings, these estimates of reliability were corrected by application of the Spearman-Brown formula to what they would be if the average number of supervisory ratings were equal to the average number of associate ratings. The corrected estimates of reliability, (shown in Table 3), are substantially equal to those obtained on the associates' ratings.

It can be seen from Table 3 that all of the scales have acceptable reliability; and that there is no choice among the scales on this basis.

CORRELATION OF MEAN RATINGS WITH ACCIDENT-RESPONSIBILITY INDEX

Although accident data were insufficient for use as a criterion of safety for the individual, it was considered desirable to select those scales which have highest common variance with an index of accident-responsibility, if such an index were obtainable on a sizable portion of the population used. It was found that 28 percent of the 189 drivers on whom sufficient data were at hand (ratings, estimates of mileage, etc.) had accidents on record. An accident-responsibility index was computed on the basis of the accident records of each driver by the following formula:

$$A.R.I. = 1000 \frac{\sum A_i R_i}{M}$$

where A_i = accidents sustained while in present assignment

R_i = responsibility for A_i as estimated from the records on a 5-pt. scale (1 = no resp., 5 = totally resp.)[1]

M = estimated number of miles driven while in present assignment

[1] - These estimates were made by the ratees' commanding officers, personnel officers, or members of their staffs. Persons making these estimates were not included among those asked to make ratings on the scales.

The distribution of this index is shown in Table 4.

TABLE 4.

DISTRIBUTION OF ACCIDENT-RESPONSIBILITY INDEX

ARI	f
400-439	2
360-399	0
320-359	1
280-319	1
240-279	1
200-239	0
160-199	1
120-159	1
80-119	4
40-79	13
0-39	29
00	136
	189

TABLE 5

CORRELATIONS BETWEEN MEAN RATINGS AND ACCIDENT-RESPONSIBILITY INDEX

Scale	Supervisors' N = 181 Ratees	Associates' N = 189 Ratees
1	.27	.18
2	.31	.06
3	.23	.14
4	.21	.10
5	.24	.10
6	.20	.10
7	.19	.08
8	.15	.04
9	.23	.08
10	.10	-.02
11	.16	.08

Table 5 shows the correlations between mean ratings and this accident responsibility index. These r's were expected to be low because of the undoubtedly low reliability of this index. However, this afforded a useful comparison of the scales. It will be noted that the supervisors' ratings have consistently higher correlations with this index than do the associates' ratings. Moreover, the latter show essentially zeros on all the scales, except, perhaps, on Scales 1 and 3.[2] It is also important to note that the lowest correlations between supervisors' ratings and this index are on Scales 8, 10, and 11. The degree to which any of these correlations might have been affected by raters' knowledge of ratees' accidents cannot be estimated.

2/ - This apparent discrepancy is considered below under Selection of Raters for Final Criterion Ratings.

INTERCORRELATION AMONG THE SCALES AND FACTOR ANALYSIS OF THE INTERCORRELATION MATRICES

The final guide for the selection of the scales for the final criterion instrument was a factor analysis of the intercorrelation matrices of the mean ratings shown in Tables 6A and 6B. The Thurstone Centroid method was used and rotation effected to the best-fitting orthogonal solution. Table 7 shows the loadings on the three orthogonal factors for each matrix.

TABLE 6A

INTERCORRELATIONS AMONG MEAN RATINGS BY SUPERVISORS ON 11 DRIVERS RATING SCALES (N = 181 RATEES)

Scales	1	2	3	4	5	6	7	8	9	10
1										
2	.78									
3	.65	.73								
4	.78	.73	.70							
5	.84	.79	.71	.86						
6	.81	.73	.63	.85	.85					
7	.73	.62	.60	.77	.80	.86				
8	.45	.44	.46	.46	.47	.50	.46			
9	.71	.78	.67	.76	.76	.69	.59	.43		
10	.46	.49	.46	.47	.50	.49	.44	.63	.51	
11	.65	.58	.44	.62	.67	.67	.67	.52	.60	.66

TABLE 6B

INTERCORRELATIONS AMONG MEAN RATINGS BY ASSOCIATES ON 11 DRIVER RATING SCALES (N = 189 RATEES)

Scales	1	2	3	4	5	6	7	8	9	10
1										
2	.81									
3	.81	.82								
4	.81	.84	.78							
5	.84	.83	.79	.86						
6	.84	.79	.75	.88	.90					
7	.79	.77	.75	.85	.89	.90				
8	.44	.54	.61	.57	.57	.58	.57			
9	.76	.83	.78	.83	.86	.82	.84	.65		
10	.56	.70	.61	.66	.71	.67	.67	.63	.74	
11	.69	.72	.67	.75	.80	.80	.79	.61	.79	.77

TABLE 7

FACTOR LOADINGS ON ORTHOGONAL AXES DERIVED FROM TWO INTERCOORELATION MATRICES OF 11 DRIVER RATING SCALES

	Supervisors' Matrix					Associates' Matrix			
	Factor					Factor			
Scale	I	II	III	h2/	Scale	I	II	III	h2/
1	.80	.30	.03	.80	1	.90	.20	-.07	.87
2	.81	.04	.13	.77	2	.82	.23	.21	.84
3	.76	.01	.15	.70	3	.85	.14	.22	.84
4	.88	.16	-.01	.85	4	.78	.44	.04	.87
5	.86	.26	.02	.89	5	.77	.51	.02	.90
6	.80	.44	-.05	.90	6	.74	.57	-.04	.92
7	.68	.54	-.01	.82	7	.71	.57	-.02	.88
8	.26	.42	.50	.60	8	.37	.60	.25	.62
9	.74	.09	.24	.74	9	.68	.51	.22	.85
10	.22	.51	.53	.69	10	.46	.45	.51	.77
11	.55	.16	.44	.69	11	.62	.33	.46	.80

Factor I appears to be similar in both matrices. Since it has highest loadings on the scales that deal with overt driving behavior and low loadings on the nondriving-performance scales (Nos. 8 and 10), we might designate this factor "rated general driving performance." At any rate this is the factor which represents the variance of most of the scales and has low loadings on the suppressor scales (Nos. 8 and 10).

Factor III is also similar in both matrices. In the associates' matrix it has loadings on Scale 10 (Appearance and Military Bearing) and Scale 11 (Maintenance). In the supervisors' matrix it also has a loading of 0.50 on Scale 8 (Like-Dislike). This seems to be an appearance factor, of both the driver and his vehicle, and it would also seem that supervisors like a driver who presents a good appearance.

Factor II is somewhat more difficult to interpret. Moreover, it does not have high loadings on quite the same scales in both matrices. It might represent some aspect of interpersonal relationships or attitudes that affect the ratings but are not correlated with overt driving behavior, at least not in the case of the supervisors.

Scales were to be selected which have high loadings on the first factor and low loadings on the other two factors in both matrices.

SELECTION OF THE FINAL SCALES

First, the six scales with highest loadings on Factor I were selected. From among these six, four were chosen on the basis of high correlation with the accident-responsibility index and low loadings on the other two factors. The scales selected were Scale 1 (Near Accidents), Scale 2 (Reaction to Sudden Changes), Scale 3 (Effect of Temper on Driving), and Scale 4 (Knowledge of Own Limitations). The distributions of mean ratings on these scales are shown in Figure 2 for associates, and Figure 3 for supervisors. It is noteworthy that the distributions of mean ratings on the scales given by supervisors have considerably greater range than the distributions of ratings given by associates.

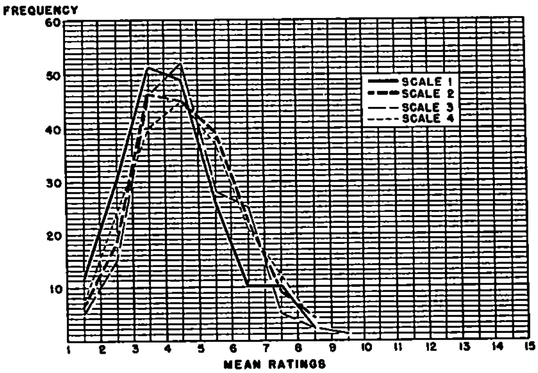

Figure 2. Distributions of mean ratings of 189 drivers rated by associates.

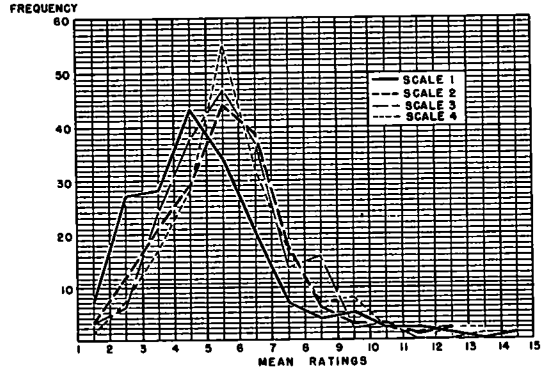

Figure 3. Distributions of mean ratings of 181 drivers rated by supervisors.

Scale 10 (Appearance and Military Bearing) was selected to precede these four scales in the final booklet. It is not scored but is included to draw off the personal feelings of the rater toward the ratee. This scale has low loadings on Factor I and high loadings on Factor III in both matrices. Also, it is less obvious than Scale 8 (Like-Dislike) which was also designed to be used as a suppressor scale.

CONSTRUCTION OF EXPERIMENTAL ITEMS FOR THE DRIVING-HABIT CHECK LIST

The driving-habit check list was developed concurrently with the scales. Suggestions for driving habits which might be considered to be associated with accident causation were sought in the pertinent literature and in consultation with safety personnel, both in the army and in civilian life; 105 statements of such driving habits were devised. These statements were reviewed with the army safety engineer, the director of the Pentagon Motor Pool, and several army drivers in order to assure their clarity, specificity, and the appropriateness of their language to army drivers.

SELECTION OF THE FINAL ITEMS FOR THE CHECK LIST

For an item to serve its purpose in the criterion instrument it would have to be (1) ratable (that is, a statement of observable behavior on which supervisors and associates could rate a driver) and (2) important to safe driving. A study was conducted, therefore, to identify those experimental items considered most important and most ratable by both drivers and supervisors in the army situation.

The 105 statements were printed in booklet form in two parts. Part I, designed to measure ratability (observability), required the subject (consultant) to check one response for each statement: (1) I know I can rate on this; (2) I would rate on this but I would not be too sure about it; or (3) I could not rate on this.

Part II, designed to measure the importance of each habit to safe driving, required the subject (consultant) to check for each statement either: (1) very important, (2) important, or (3) not important.

This booklet was administered to the same drivers and supervisors who participated in the rating sessions described earlier in this report (158 drivers and 35 supervisors).

On 21 of the items the responses "I know I can rate on this" and "very important" were checked by at least 49 percent of the drivers and at least 49 percent of the supervisors. These items and the percentages of judgments are shown in Table 8. In order to select the 15 items for the criterion instrument from these 21, six of the items (Nos. 9, 25, 36, 48, 75, and 84) were eliminated on the basis of the judgments of the investigators.

TABLE 8

21 BEHAVIOR STATEMENTS RATED "VERY IMPORTANT" AND "I KNOW I CAN RATE ON THIS" BY 49 PERCENT OR MORE OF DRIVERS AND SUPERVISORS

Item No.	Behavior Statement	Percent Checking "Very Important" A*	S*	"I Know I Can Rate on This" A*	S*
1.	Breaks the speed limit	59	66	52	60
2.	Drives too fast for road conditions	68	89	51	69
3.	Doesn't stay on his side of the road	73	91	54	49
4.	Ignores stop lights or signs	77	89	61	63
5.	Doesn't give the right of way to other drivers	65	60	58	60
6.	Passes on curves and hills	76	83	49	51
7.	Doesn't signal for stops or turns in advance	57	71	60	63
8.	Doesn't check brakes before driving	71	77	51	51
9.	Doesn't slow down at intersections when he has the right of way	52	63	57	60
14.	Follows other vehicles too closely	59	77	60	69
20.	Pulls away from the curb without looking back for oncoming traffic	66	80	51	51
23.	Takes chances when driving	50	71	51	54
25.	Gets into accidents with other vehicles	64	63	50	74
36.	"Horses around" when he's driving	62	80	50	54
37.	Shows off when driving	61	74	49	60
44.	Doesn't cut wheels to curb when parking on a hill	66	80	49	49
45.	Backs up without looking behind	72	80	53	57
48.	Drives with dirty windshield	55	57	58	71
75.	Swings too wide on turns	55	60	50	54
84.	Drives faster than the other traffic	54	49	51	51
92.	Fails to turn in his vehicle for repairs promptly	60	60	50	77

*A = Associates
*S = Supervisors

RELATION BETWEEN SELECTED CHECK LIST ITEMS AND EXPERIMENTAL RATING SCALES

In order to obtain some insight into the relationship between the continua of the rating scales and these 15 items, eight staff members of the Personnel Research Section were asked to classify the 15 items, following these instructions: "Indicate under which of the rating scales you would subsume each of the driving habits if you were doing the rating."

No item was to be assigned by the same PRS judge to more than three different scales. A shortcoming of this method is that primary, secondary, and tertiary assignments receive the same weight. However, the results, shown in Table 9, indicate such a high predominance of assignments to Scales 5 and 6 that the findings are readily interpretable.

TABLE 9

ASSIGNMENTS OF SELECTED CHECK LIST ITEMS TO RATING SCALE CONTINUA

Scale Number	Item Number 1	2	3	4	5	6	7	8	14	20	23	37	44	45	92
1					1		1	1	1	1	2	1			
2				1					1		1				
3					2										
4				1							1				
5	7	6	8	3	5	8	7	5	6	8	6	6	5	8	1
6	8	6	3	8	5	5	6	3	5	5	7	7	5	5	3
7	1	1	1	3	1	1	2			2		3			
8					1										
9		2	2						3				2	1	
10															
11			1					5					2		8

FINAL CRITERION INSTRUMENT

The final criterion instrument consists of: (1) the practice rating scale; (2) the suppressor scale (not to be scored); (3) the four criterion scales; and (4) the driving-habit check list.

SCORING OF THE FINAL CRITERION INSTRUMENT

A driver's score on the scales is the mean of ratings received on the four criterion scales, supervisors' and associates' ratings taken together (the basis for this decision is furnished below, under "Selection of Raters"). His score on the check list is the mean number of checks received. On the basis of the judgments of the investigators with respect to the relative variance contributed by each of these two measures, a weight of 2 for the mean rating and a weight of 1 for the mean check list score was considered to yield about the optimum composite criterion score.

SELECTION OF RATERS FOR FINAL CRITERION RATINGS

Tied in with the problem of selection of scales was the problem of selection of raters. It is evident from the larger means and standard deviations of the mean ratings (Table 3) and from the consistently higher correlations with the accident-responsibility index (Table 5), that the supervisors' ratings are superior for our purpose. But, since it is often impossible to obtain ratings from more than two or three supervisors and this criterion is postulated upon having several ratings on each driver, the possibility of supplementing these with the ratings of selected associates was explored.

The first attempt to select among associate raters was made on the basis of grade. The correlations of different grades of associates' ratings with supervisors' ratings are shown in Table 10. These correlations furnished no basis for selection among associate raters.

TABLE 10

CORRELATIONS BETWEEN RATINGS BY SUPERVISORS AND RATINGS BY DIFFERENT GRADES OF ASSOCIATES

Scale	Ratings by Sgts.	Ratings by Cpls.	Ratings by Pfcs and Pvts.	Ratings by all Associates
1	0.42	0.37	0.34	0.37
2	0.42	0.36	0.33	0.35
3	0.40	0.51	0.38	0.42
4	0.35	0.48	0.45	0.43
5	0.52	0.52	0.42	0.47
6	0.53	0.49	0.41	0.46
7	0.51	0.53	0.32	0.43
8	0.18	0.23	0.15	0.18
9	0.42	0.45	0.37	0.41
10	0.59	0.44	0.38	0.44
11	0.41	0.44	0.32	0.37

In the light of another study by the Personnel Research Section (7), it was indicated that GCT level of raters is more highly related to validity of ratings than is grade. Since the raters used in the present study were identifiable only by grade and group, two groups of corporal raters were selected for comparison, Group A being composed mostly of corporals with GCT scores below 90 and Group B being composed mostly of corporals with GCT scores above 90.3/

3/ - The exact mean scores were unobtainable. Group A actually consisted of 33 of 40 corporals whose GCT distribution was known, and Group B consisted of 17 of 24 corporals whose GCT distribution was known. In making the comparisons reported here, the assumption was made that the GCT scores of those--seven in each group--who did not participate in the rating sessions were more or less randomly distributed.

The correlations between accident-responsibility index and the ratings accomplished by these two groups of corporals are shown in Table 11. Although comparison on any single scale may be inconclusive, the pattern of correlations favors the higher level group with marked consistency. On this basis, the decision was made to obtain criterion ratings from supervisors and only those associates whose GCT scores were at least 90.

TABLE 11

CORRELATIONS BETWEEN MEAN RATINGS AND
ACCIDENT-RESPONSIBILITY INDEX: TWO GROUPS OF CORPORAL RATERS

Scale	Group A* Mean GCT Below 90	Group B* Mean GCT Above 90
1	0.24	0.38
2	0.15	0.28
3	0.22	0.32
4	0.22	0.31
5	0.15	0.30
6	0.18	0.38
7	0.08	0.27
8	0.11	0.17
9	0.15	0.35
10	0.14	0.15
11	0.15	0.27

*In Group A, 33 corporals rated 78 drivers.
In Group B, 17 corporals rated 34 drivers.

SUMMARY

Accident records and road tests were considered and abandoned as bases for a criterion of safe driving in the army situation. Observations and judgments of drivers, supervisors, and associate drivers were then considered and an instrument, including rating scales and a check list, was developed for their quantification. A population of 189 drivers was rated on 11 experimental scales by an average of 4.8 supervisors and 12.5 associates. Of the 11, four scales were finally chosen on the bases of (1) reliabilities, (2) correlations with an accident-responsibility index, (3) intercorrelations among the scales, and (4) results of a factor analysis of these intercorrelations. The same raters were asked to indicate, for each of 105 descriptions of unsafe driving habits, how ratable (observable) the behavior is and how important it is to safe driving. The 15 statements adjudged most ratable and most important were selected for the final check list. The mean rating on the four scales receives a double weight and the mean number of checks received has unit weight in the composite criterion score.

REFERENCES

1. Bransford, T. Relation of performance on driver's tests to automobile accidents and violations of traffic regulations in the District of Columbia. Unpublished Ph.D. dissertation, Amer. Univ., 1939.

2. Brown, C. W. and Ghiselli, E. E. Accident proneness among streetcar motormen and motor coach operations, J. Appl. Psychol., 1948, 32 (1).

3. Farmer, E. and Chambers, E. G. A Study of Accident Proneness among Motor Drivers. Report No. 84, Medical Research Council, Industrial Health Research Board, London: His Majesty's Stationery Office, 1939.

4. Horst, P. A generalized expression for the reliability of measures. Psychometrika, 1949, 14(1), 21-31.

5. Slocombe, C. S. Consistency of operating efficiency. Person. J., 1930, 8, 413-414

6. Staff, Personnel Research Section, The Adjutant General's Office. Personnel research in the Army, VI. The selection of truck drivers. Psychol. Bull., 1943, 40(7).

7. PRS Report on PJ 4903-09, Unpublished. Effectiveness of Ratings Accomplished by Raters with Aptitude Area I Scores Below 90. Personnel Research Section. The adjutant General's Office, Department of the Army.

8. The Accident Prone Driver. House Document No. 462, Part 6. Bureau of Public Roads. Washington: Government Printing Office, 1938.

RELATION between PSYCHOLOGICAL TESTS and DRIVER PERFORMANCE

D. J. Moffie, Andrew Symmes, and Charles R. Milton
North Carolina State College

THE North Carolina State College through its Extension Division established a Driver Training School on August 8, 1949 for training tractor-trailer truck drivers.

Psychological research was one of the aims of the training program when the school was organized. The research program embodied interviews, psychological testing, psychological paper-and-pencil tests, and accident follow-up. The purpose of this report is to show the relationship between psychological tests and tractor-trailer driver performance.

Using the Cornell Word Form Test, the Eno Foundation for Highway Safety (3) found that accident repeaters were more emotionally unstable than the accident-free drivers. The drivers were commercial-vehicle operators, as well as noncommercial, and were matched (N - 193) as to sex, type vehicle operated, and mileage. The difference found between the two groups was significant at the 5-percent level of confidence.

Ghiselli and Brown (2) compared accident records for the first 5 weeks of employment with psychological test scores for 67 taxicab drivers. With accidents, a mechanical-principals test correlated 0.11; a total-interest score, 0.23; dealing with people, 0.11; and related occupations, 0.20.

Ghiselli and Brown (1) found that a learning period exists for street-car and motor-coach operators. Improvements in accident records were evident after the first 6 or 7 months of operation, and accident rates were reduced considerably after this time. They found that the reduction of accidents approached the shape of the common-learning curve. It was concluded that "for motor-coach and street-car operators it is simple enough to learn the motor coordinations involved, but complex activities such as judgments of speed, spatial relations, and division of attention, performed under conditions of stress are the difficult aspects of learning the job."

Research by The Adjutant General's Office (4) on selection of truck drivers yielded low coefficients of reliability for criterion data. One index of driving ability was the score the driver obtained on the Road Test Check List. An examiner observed and checked items best describing the driver's performance in a standardized skill test situation. The reliability coefficient of the check list was 0.82 by the split-half method after correction by the Spearman-Brown formula. Scores on the check list given by different examiners at different times correlated 0.53. General ratings of 127 men by several examiners on the same day yielded an average correlation of 0.72 between judgments of the different raters. The article states that the low reliability of these criteria could very likely be improved by proper training of examiners.

METHODS AND PROCEDURES OF THIS STUDY

The paper-and-pencil tests used were the Otis SA Test of Mental Ability (Higher Form B), the Bennett Test of Mechanical Comprehension (Form BB), the Kuder Vocational Interest Test (Form B), the Bernreuter Personality Inventory, and the Minnesota Multiphasic Personality Inventory (Group Form). All tests were administered during the trainees' 6-week training period.

The subjects are over-the-road drivers (haul freight over long distances as opposed to the local-delivery personnel) and were of the dual-operation category (two drivers ride together, one drives and the other sleeps). All the drivers were from six large companies operating on the East Coast with home terminals in North Carolina.

Accident records were found to be the most suitable index for the criterion of job performance. The accident data were secured by the writers who were granted access to the files of each company employing drivers from the school. The accident records were found to be adequately maintained and were labeled in a similar manner. All companies classified accidents as preventable or nonpreventable. In some cases the chargeability of the accidents were determined by a board of safety supervisors in the personnel and safety departments. However, in some companies classifications were made by the safety director alone. An accident was preventable if the driver could have foreseen and avoided the accident by any conceivable means and nonpreventable if the accident could not have been prevented.

An accident score was derived from each man's record in the following manner:

$$\frac{\text{Months Worked}}{3 \times \text{no. preventable accidents} + 2 \times \text{no. of nonpreventable accidents.}}$$

Rather than use the number of accidents alone as the criterion, the writers thought that some adjustment should be made for experience. The longer a man has driven the more susceptible he had been to hazardous conditions and accident-producing situations. These factors fail to remain equivalent if the drivers have been operating for varying lengths of time. By utilizing months worked in the above formula, such inequalities have been considerably alleviated, although not perfectly controlled.

After going over approximately 400 accident records, the writers also felt that some distinction should be made between the accident categories of preventable and nonpreventable in arriving at an accident score. In almost every nonpreventable accident examined, there were factors operating or conditions present which lessened or prevented the possibility of a driver foreseeing and avoiding an accident. Therefore, it was felt that a system of weights would tend to equate variation between these two types of accidents. Figure 1 shows that the distributions of preventable and nonpreventable accidents are practically identical. The variances of the two curves are not significantly different from each other. This justifies the addition of the two scores.

One other factor considered in the analysis of the group should be mentioned: No driver was included in the study who had not driven at least 6 months. This would represent approximately 30,000 mi. of driving. A follow-up of the school's trainees has revealed that of the number who have accidents (some had no accidents), the greatest proportion had the majority of their accidents during the first 6 months.

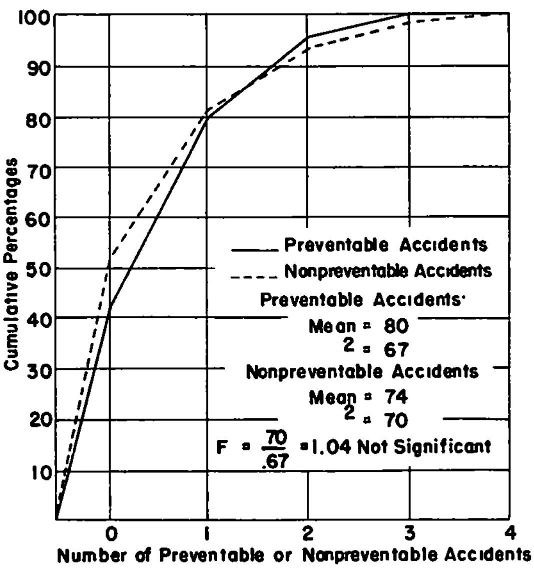

Figure 1. Distributions of cumulative percentages of preventable and nonpreventable accidents.

The distribution of accident scores for the 6-to-18 months group as shown in Figure 2 has a tendency towards bimodality with the majority of the subjects having low accident scores. Skewness is positive. It is realized that, in some studies which have utilized accidents alone, the distribution has been a poisson one.

Although only 31 drivers were available who had worked more than 12 months, bimodality found in the 6-month group disappeared.

The study is divided into three parts: In Part 1 an accident group of 30 drivers was matched with an accident-free group on the basis of months worked. These groups were then compared on the basis of differences in mean test scores and t ratios were utilized to test the significance of the difference. The accident-free group had a mean of 9.34 months of work and a sigma of 3.18; the accident group had worked 9.36 months and had a sigma of 3.21.

In Part 2 the group consisted of approximately 100 subjects who had driven an average of 10.82 months with a sigma of 3.65. Pearson-product moment correlations were utilized to determine the relationship between the psychological tests and the criterion.

Part 3 consisted of approximately 31 drivers who had been on the job from 12 to 18 months. The average time driven was 15.72 months with a sigma of 1.84. A Pearson product-moment correlation was used to determine the relationship between test scores and the criterion.

RESULTS

The t ratios obtained between the matched groups and the correlation coefficients for the two groups working varying lengths of time are shown in Tables 1 and 2, respectively.

The direction of relationship for the matched groups can be determined by examining the mean scores of the two groups.

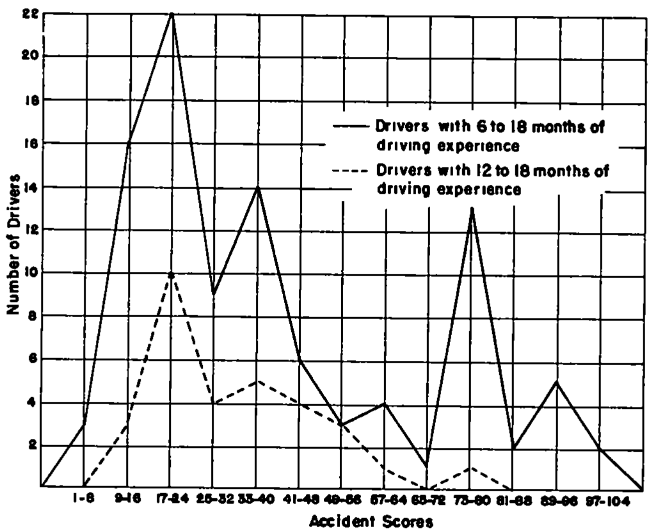

Figure 2. Distribution of accident scores for truck drivers.

In the first part of this study, significant differences were found on the B1-N, B2-S, and B4-D scales of the Bernreuter Personality Inventory, the Bennett Test of Mechanical Comprehension, and the computational scale of the Kuder Vocational Interest. On the Bernreuter scales, the accident-free group tended to make higher scores on B1-N and lower scores on B2-S and B4-D than the accident group. This means that the accident-free drivers tend to be tense, less self-sufficient, and less dominant than the accident driver. On the mechanical-comprehension tests, the accident-free driver made higher scores indicating great comprehension of mechanical principles. This was not true, however, on computational interest—accident drivers have higher interest in computational activities.

In the second part of the study, a significant negative correlation was obtained between the B2-S scale, and the accident index; a significant positive correlation was obtained between the Bennett test and the criterion. These results tend to substantiate the results disclosed in Part 1.

In the third part, only one correlation of significance was obtained and this was between the Kuder Mechanical Interest scale and the criterion indicating positive relationship.

Differences in results among the three groups were indicated. The writers are inclined to believe, however, that such differences may be due primarily to sampling and operation of uncontrollable factors. There may be changes in psychological composition due to the hazardous conditions of the job, the dual type of operation, and the disruption of home life. Also, selective factors, such as discharge from the job and voluntary terminations, may alter the nature of the group so that the psychological traits of drivers who remain on the job longer vary appreciably from the ones who stay for shorter intervals. In this connection it should be mentioned that ideally one should give these tests periodically to detect and evaluate such personality changes.

None of the tests show significant relationships with the criterion for all three groups. However, the Bennett Test of Mechanical Comprehension and the B2-S scale of the Bernreuter gave results significant at the

TABLE 1

MEANS AND STANDARD DEVIATIONS FOR PSYCHOLOGICAL TESTS FOR ACCIDENT AND ACCIDENT-FREE GROUPS AND RELATIONSHIP OF THE TWO GROUPS ON THESE TESTS EXPRESSED BY t RATIOS

Psychological Test	Accident Group N = 30		Accident-free Group N = 30		
	Mean	σ	Mean	σ	t
Otis Self Administering Test	24.53	8.50	26.10	9.08	.584
Bennett Mechanical Comprehension	21.47	8.67	27.90	10.51	2.568[a]
Kuder - Mechanical	95.13	13.32	96.50	11.63	.411
Computational	34.54	7.20	30.60	6.26	2.155[a]
Scientific	63.93	11.37	62.30	13.53	.467
Persuasive	67.33	15.14	67.67	12.76	.104
Artistic	47.47	12.07	48.70	11.54	.390
Literary	38.20	9.57	38.03	7.61	.072
Musical	13.07	8.38	14.83	7.06	.779
Social Service	77.67	14.57	78.20	14.60	.147
Clerical	54.60	10.49	48.87	8.70	1.647
Bernreuter - B1-N	19.43	17.90	33.47	22.36	3.146[b]
B2-S	42.43	23.58	31.20	19.15	2.138[a]
B4-D	67.43	19.61	53.83	23.19	2.455[a]
M.M.P.I. - K Scale	17.30	4.30	16.30	4.67	.908
Hypochondriasis	16.37	4.56	17.53	3.98	1.169
Depression	16.77	2.81	16.77	3.35	.000
Hysteria	1.63	1.35	2.57	2.70	1.800
Psychopathic Deviate	16.30	2.87	16.57	3.60	.312
Interest Scale	20.60	4.10	21.30	4.86	.716
Paranoia	7.47	2.16	8.13	2.31	.803
Psychasthenia	7.03	5.36	7.53	6.08	.408
Schizophrenia	7.00	5.14	8.17	4.80	.742
Hypomania	17.00	3.48	16.40	3.02	.648

a/ - Significant at 5-percent level.
b/ - Significant at 1-percent level.

5-percent level for the drivers who had worked from 6 to 18 months and also for the matched groups. In general one may say that the more proficient drivers have a better understanding of general mechanical principles than the ones who have poorer accident records. One may also surmise that the negative relationship between self-sufficiency and the better drivers indicates that they are dependent upon others and therefore more hesitant and cautious in their driving habits. These drivers are more likely to size up traffic situations before proceeding.

However, one cannot select a driver on the basis of a particular area of interest or one personality trait without examining all test results. It is the pattern of abilities, interests, or personality traits that should receive consideration collectively rather than any one or two taken alone. In this respect a multiple correlation would be the best statistical treatment for one to utilize because it permits a multiple analysis of numerous factors and their interaction.

TABLE 2

MEANS AND STANDARD DEVIATIONS FOR PSYCHOLOGICAL TESTS AND PEARSON PRODUCT MOMENT CORRELATIONS FOR EACH TEST WITH ACCIDENT SCORES FOR TWO GROUPS OF DRIVERS

Psychological Test	Worked from 6 to 18 months				Worked from 12 to 18 months			
	Mean	σ	N	r	Mean	σ	N	r
Otis Self-Administering	25.18	9.05	99	-.040	25.70	9.70	30	-.255
Bennett Mech. Compr.	24.64	10.09	99	.197[a]	26.80	10.21	30	.185
Kuder Mechanical	96.40	12.96	100	.062	100.64	5.96	31	.443[b]
Computational	31.52	7.20	100	-.108	29.48	6.31	31	-.239
Scientific	62.41	13.14	100	-.039	62.03	14.27	31	-.087
Persuasive	68.53	14.27	100	.058	69.68	13.80	31	.164
Artistic	47.90	11.18	100	.058	45.29	10.78	31	-.142
Literary	39.22	10.61	100	-.027	41.23	11.14	31	.074
Musical	15.34	8.19	100	-.004	17.10	9.22	31	-.181
Social Service	76.76	14.26	100	.056	75.52	14.77	31	.165
Clerical	50.97	12.50	100	-.054	49.19	9.37	31	.181
Bernreuter B1-N	26.82	21.71	98	.155	28.28	20.37	29	-.080
B2-S	35.30	22.37	98	-.228[a]	34.00	21.75	29	-.147
B4-D	57.06	23.24	98	.033	55.79	22.59	29	.092
MMPI K Scale	16.31	5.21	100	.022	14.35	4.01	31	.017
Hypochondriasis	16.29	2.47	100	.064	15.29	2.25	31	.010
Depression	16.24	4.44	100	-.034	15.68	3.61	31	-.202
Hysteria	2.06	4.19	100	.069	2.03	4.42	31	.148
Psychopathic Deviate	16.58	4.52	100	.040	15.10	3.82	31	-.034
Interest Scale	20.49	4.15	100	.027	21.23	3.78	31	.028
Paranoia	7.89	2.87	100	-.007	8.19	3.20	31	.106
Psychasthenia	7.10	5.32	100	-.082	8.35	5.42	31	-.268
Schizophrenia	7.03	5.48	100	-.104	8.77	6.88	31	-.188
Hypomania	17.22	4.27	100	-.075	18.32	4.61	31	.070

a/ - Significant at 5-percent level.
b/ - Significant at 1-percent level.

CONCLUSIONS

This study, like many in the past, has shown some relationships between psychological traits and driver performance. Unlike many, it has disclosed the importance of the personality of the driver as a factor in safety. Conclusions are as follows:

1. Accident rates, even though determined on the basis of miles driven per accident, continue to follow the Poisson distribution. This implies that correlational indices need to be interpreted with caution.

2. Safe drivers tend to be more tense, less self-sufficient, less dominant, as shown by the Bernreuter Personality Inventory. These same trends are demonstrated on the MMPI test (hypochondriasis and hysteria scale) even though differences were not shown to be statistically significant.

3. Safe drivers tend to make higher scores on the Bennett Mechanical Comprehension test.

4. The first 6 months of driving experience appears to be a learning period because of the high incidence of accident rates.

TABLE 3

COMPARISON OF THREE GROUPS OF DRIVERS, ACCIDENT VERSUS NONACCIDENT, THOSE WHO WORKED 6 TO 18 MONTHS, AND THOSE WHO WORKED 12 TO 18 MONTHS

Psychological Test	Accident vs. Nonaccident t ratio	Accident vs. Nonaccident Direction	Worked 6 to 18 Months r	Worked 12 to 18 Months r
Otis Self Administering	.584	N-A	-.040	-.255
Bennett Mech. Compr.	2.568a/	N-A	.197a/	.185
Kuder Mechanical	.411	N-A	.062	.443b/
Computational	2.155a/	A	-.108	-.239
Scientific	.467	A	-.039	-.087
Persuasive	.104	N-A	.058	.164
Artistic	.390	N-A	.058	-.142
Literary	.072	A	-.027	.074
Musical	.779	N-A	-.004	-.181
Social Service	.147	N-A	.056	.165
Clerical	1.647	A	-.054	.181
Bernreuter B1-N	3.146b/	N-A	.155	-.080
B2-S	2.138a/	A	-.228a/	-.147
B4-D	2.455a/	A	.033	.092
MMPI K Scale	.908	A	.022	.017
Hypochondriasis	1.169	N-A	.064	.010
Depression	.000	-	-.034	-.202
Hysteria	1.800	N-A	.069	.148
Psychopathic Deviate	.312	N-A	.040	-.034
Interest Scale	.716	N-A	.027	.028
Paranoia	.803	N-A	-.007	.106
Psychasthenia	.408	N-A	-.082	-.268
Schizophrenia	.742	N-A	-.104	-.188
Hypomania	.648	A	-.075	.070

a/ - Significant at 5-percent level
b/ - Significant at 1-percent level

REFERENCES

1. Ghiselli, E. E., and Brown, C. W. Learning in Accident Reduction. J. Appl. Psychol., 1947, 31, 580-582.

2. Ghiselli, E. E., and Brown, C. W. The Prediction of Accidents of Taxicab Drivers. J. Appl. Psychol., 1948, 33, 540-546.

3. Personal Characteristics of Traffic Accident Repeaters. Eno Foundation for Highway Traffic Control. Saugatuck, Connecticut, 1948.

4. Personnel Research in the Army. VI. The Selection of Truck Drivers: By Staff, Personnel Research Section, Classification and Replacement Branch, The Adjutant General's Office. Psychological Bulletin, 40, 1943.

AGE and SEX in RELATION to ACCIDENTS

A. R. Lauer
Director of the Driving Research Laboratory
Iowa State College

A PROPER analysis of any problem is essential to an adequate solution. It seems that the problems of age and sex in relation to automobile and traffic accidents have been grossly confused, due partly to the emotionality of writers on the subject.

The two are inseparable, since vast differences between the sexes are noted at various age levels. It is unintelligible to speak of the problems concerning one without a proper clarification of the limits of the other.

In order to treat an ailment, the doctor must find the origin of it or the part affected. Medication for eczema is not likely to help one with arthritis. Too often we try to reduce accidents by shot-gun remedies. Until the causes and those responsible for most of the trouble are singled out the proper methods of accident reduction cannot be determined.

The purpose of this paper is primarily to untangle some of the faulty conclusions that have been widely discussed about youthful drivers, women drivers, and older drivers. It is also hoped that it may stimulate greater care in coding the exact ages of drivers by the state licensing bureaus. Unfortunately, huge sums of money are spent annually in hoarding accident records and making compilations of simple tables, but little if any really scientific analysis is made of them. It is time that this practice is changed. We need to digest the results of accident reporting so that leads may be obtained which will be useful in educational,legislative, and enforcement circles to aid in accident reduction.

A very important aspect of the problem is the uneven grouping of ages for comparative purposes. The terminologies used are even confusing. Writers of popular articles on "teen-age" drivers sometimes construe the term to include every age up to 25 years. One state was known to report that 3 percent of its licensed drivers were 20 and below and were having half the fatal accidents from 12 midnight until 6 a.m. A recheck of the files showed that nearly 8 percent of their licensees were 20 and below, since 20-year olds were classified with "teen-agers."

It was with the hope that some of these confusing issues might be cleared up that a 5-year study of age and sex has been undertaken at Iowa State College. The specific hypothesis set up as a basis of the present paper may be stated as follows: Reported accidents are distributed evenly throughout the driving population according to density of population, age, and number of licensees. A corollary hypothesis to be tested may be formulated as: Accidents are distributed evenly throughout the driving population on the basis of miles driven by each age group.

In addition, answers to the following questions are sought through various studies in progress: Are age trends in accident involvement,as secured from the accident files, consistent? Are accident reports in a given community an index to its fatality record? Is reporting uniformly carried out within a state and is it in anyway associated with density of population? If the hypotheses set up are confirmed, what would be the most fruitful immediate approach to the motor-vehicle accident problem? If the hypotheses are rejected, where should we start to improve drivers?

It is obvious that, as a rule, shot-gun methods in education, engineering, or any other field of endeavor will be less effective than systematically planned and organized effort. The least that can be left to chance the better. It is assumed that accident reporting is impartially carried out with respect to sex and age groups within a given municipality, state or territory.

METHOD AND PROCEDURE

This study was undertaken in cooperation with the Department of Public Safety in Iowa, the data presented being from Iowa. Continuation studies in cooperation with the motor-vehicle departments of Illinois, Minnesota, and Pennsylvania are now in progress and will be presented at a later time.

The method is of statistical and analytical nature and consists of drawing a systematic sampling of the drivers' license files in the state by examining and tabulating the data from every two-hundredth card of the 1,300,000 licenses in the state. The cards are jacketed and contain information on all accidents, violations, and revocations made by the licensee since the drivers license law became effective in Iowa in 1933. Overload and nonmoving violations were not included, since they may not be the fault of the driver. It is the information of the writer that standing violations, such as parking tickets, are not entered in the records.

It suffices to say that the reporting system developed gradually, as in other states, and that reporting became more efficient as the patrol was increased and more attention was given to accident reporting by the local traffic courts and enforcement agencies. Consequently, the increased number of accidents reported does not necessarily serve as a true index of the total number of accidents occurring within a given period. This is shown by the sharp jump in 1948 when the financial-responsibility law went into effect, as shown in Table 3.

After the list of 7,692 cases comprising the sample of licenses used in this study was drawn and prepared, a double, specially printed postcard was sent to each person requesting further information regarding the mileage driven annually by day and night. Other questions were asked about the way the person learned to drive, mileage on the car, age of car, model, ownership of car, and certain other related data. Only items referring to age, licensure, mileage, violations, and accidents will be reported here as taken from the state records.

The data were assembled and punched on IBM cards for analysis by age groups, sex, and location and according to mileage driven annually. Accidents, violations, age, and personal data were taken from state files. The data were recorded by years of age to the nearest birthday, but only accidents and violations for the years 1948 and 1949 were considered in this study, due to the lack of consistency prior to that date (see Table 3.)

RESULTS

It was first considered necessary to make an estimate of the number of potential drivers in a state the size of Iowa for each age group. This is shown in Table 1 and was based upon 1940-census figures. The 1950-census data were not available at the time, and it was further considered that the percentage change would not likely be significant as it may well represent certain parameters of the population for the various age levels.

TABLE 1

STRUCTURE OF THE DRIVING POPULATION
(Estimates based on U.S. Census)

Age	Population Licensed Men	Population Licensed Women	Licensees Men	Mileage Index	Licensees Women	Mileage Index
	%	%	%		%	
15	7.7	0.9	0.234	32	0.026	10
16	20.0	3.8	0.624	42	0.117	14
17	60.9	11.7	1.651	58	0.351	16
18	67.0	25.8	2.145	73	0.845	18
19	63.5	18.2	1.885	89	0.559	20
20	54.9	17.4	1.573	99	0.533	22
21	56.1	15.7	1.651	105	0.468	23
22	48.2	26.1	1.352	114	0.767	25
23	68.6	15.1	1.991	121	0.442	27
24	76.9	17.0	2.158	127	0.494	28
25	67.8	24.8	1.924	124	0.741	29
26	69.0	23.5	1.911	131	0.676	30
27	68.7	30.8	1.859	138	0.845	31
28	71.6	29.6	1.950	136	0.845	31
29	61.2	21.7	1.612	130	0.585	32
30	72.2	22.5	2.054	129	0.676	33
31	67.5	23.5	1.586	126	0.559	33
32	67.7	19.4	1.794	123	0.533	34
33	66.4	20.9	1.586	121	0.507	34
34	67.5	19.7	1.638	118	0.468	35
35	55.7	19.4	1.430	116	0.507	36
36	55.4	23.6	1.547	114	0.559	37
37	66.9	28.7	1.638	112	0.637	37
38	60.2	27.9	1.391	111	0.676	39
39	57.7	25.6	1.404	112	0.624	40
40	54.1	28.5	1.430	109	0.663	41
41	67.2	42.9	1.261	109	0.780	41
42	57.4	20.4	1.365	108	0.481	43
43	62.7	25.0	1.326	108	0.520	44
44	60.0	26.0	1.222	108	0.520	46
45	54.4	24.6	1.274	108	0.624	47
46	60.3	19.2	1.222	108	0.364	48
47	68.6	23.9	1.417	108	0.468	49
48	65.2	21.8	1.389	109	0.429	50

TABLE 1 - continued

Age	Population Licensed Men	Women	Licensees Men	Mileage Index	Licensees Women	Mileage Index
	%	%	%		%	
49	50.7	25.4	1.001	109	0.533	50
50	55.7	26.4	1.261	108	0.494	50
51	63.6	27.4	1.066	108	0.416	50
52	52.4	21.8	1.053	109	0.403	50
53	64.5	31.5	1.092	107	0.494	49
54	79.4	37.0	1.365	106	0.598	47
55	60.1	23.4	1.014	105	0.377	46
56	62.8	20.5	0.988	104	0.299	43
57	77.5	24.6	1.092	102	0.325	40
58	65.7	12.6	0.975	100	0.182	35
59	65.3	21.2	0.845	96	0.273	32
60	67.3	21.4	0.871	92	0.286	30
65	66.9	21.4	0.728	75	0.234	23
70	57.8	23.8	0.433	71	0.182	18
75	45.7	8.8	0.208	70	0.043	15
80	65.8	4.6	0.164	69	0.001	12
85	61.4	0.0	0.060	61	---	--
90	81.4	0.0	0.020	50	---	--

Table 1 contains the fundamental data necessary to graph or otherwise convert accident frequencies into meaningful form. At times trends are real, even though statistical evaluations at any one point along the curve may not show a significant difference. Experimental or sampling errors may account for occasional deviations which, on first consideration, seem out of line, and thus false conclusions regarding trends may be drawn.

Comparison may be made from this table at any age up to 60, then by 5-year intervals up to age 90. The number of drivers in a given age group in Iowa may be closely estimated by sex, also the percentage at each age level that are licensed, and the percentage of all licenses in the state held by any age group. The stated annual mileage as obtained from a poll of the sampling drawn is given by the mileage index. This indicates the nearest number in hundreds of miles.

From such a table it is possible to group ages in any fashion desired and fairly reliable estimates may be made from the results. Unless the basic constants are approximately known, no valid conclusions can be drawn from accident figures.

Another source of error commonly found is anachronism. A certain group may be studied along with others of different ages. One individual may be noted for his accident record. It may be found that he has had six accidents. No account is made of when he had them, at what age or over what period of time. Table 2 shows data on male drivers plotted as of the age they were when the accident occurred. Conditions no doubt change. By taking all accidents occurring at a given age for several years we get a better estimate of the performance of a given age level. That the reporting index at least changes during a period of time is shown in Table 3. Thus a cross-section study over several years is probably more meaningful than any short-time study, as long as correction is made for the age at which accidents occur.

TABLE 2

REPORTED ACCIDENTS ALL AGE GROUPS[a]
(Male Drivers)

(Based on a preliminary tabulation of 1,806 accidents from cases drawn.)

Age	Percent	Age	Percent	Age	Percent	Age	Percent
13	0.0574	35	1.4917	57	0.6885	79	0.2295
14	0.0574	36	2.1801	58	1.0901	80	----
15	1.1474	37	1.8360	59	0.8032	81	0.1147
16	2.0080	38	1.9507	60	0.4590	82	0.0574
17	3.0981	39	2.0080	61	0.5164	83	0.0574
18	3.4997	40	1.5491	62	0.3442	84	----
19	4.1882	41	1.6638	63	0.8032	85	----
20	3.6145	42	1.7785	64	0.2870	86	----
21	3.6718	43	1.8360	65	0.5164	87	----
22	4.3603	44	1.2622	66	0.2870	88	----
23	3.2702	45	1.6638	67	0.2870	89	0.0574
24	3.4997	46	1.5491	68	0.4590	90	----
25	3.7866	47	0.9753	69	0.4590	91	----
26	4.1882	48	0.7458	70	0.1721	92	----
27	3.2129	49	1.3769	71	0.5164	93	----
28	3.3850	50	1.0327	72	----	94	----
29	2.2375	51	0.8032	73	0.1721	95	----
30	2.2950	52	1.2048	74	0.0574	96	----
31	2.6391	53	0.9753	75	0.0574	97	----
32	1.7785	54	1.0901	76	0.1721	98	----
33	2.0080	55	0.8606	77	0.1147	99	----
34	2.1801	56	1.0327	78	0.1721		

a/ - Represents part of sampling of every 200th card of all accidents reported to the state from 1933-1950 inclusive to March 1950. Tabulated by age at the time they were reported.

TABLE 3

ALL ACCIDENTS AS REPORTED BY YEARS IN IOWA
(Male Drivers)

Year	Percent	Year	Percent	Year	Percent
1930	------	1937	1.3196	1944	3.4997
1931	------	1938	2.3523	1945	4.3029
1932	------	1939	4.0161	1946	0.3442
1933	------	1940	5.3930	1947	8.6059
1934	------	1941	5.9667	1948	29.2025[a]
1935	------	1942	4.4177	1949	27.0787
1936	0.0574	1943	3.3276		

a/ - Financial Responsibility Law enforced after January 1, 1948.

Our first hypothesis set up for testing may be considered at this time. Since temporal distribution of reporting varies, what may be expected of spatial reporting? Within a state, is reporting quite uniformly carried out throughout the different districts, counties and population areas? A statistical evaluation of approximately equal population groups in Iowa combined according to population are presented in Table 5.

TABLE 4

DISTRIBUTION OF ACCIDENTS ACCORDING TO DENSITY OF POPULATION

Group	Number of Counties	Accidents: Actual	Accidents: Expected	X^2
I	1	220	111.5	70.336b/
II	2	124	102.9	4.317a/
III	3	140	133.2	.342
IV	5	126	127.5	.018
V	5	94	98.3	.190
VI	8	83	105.8	4.938a/
VII	10	96	109.0	1.557
VIII	10	88	100.2	1.484
IX	11	106	98.0	.646
X	12	73	95.8	5.423a/
XI	15	88	116.7	7.057a/
XII	17	77	95.9	3.736

a/ - Significant at the 5-percent level.
b/ - Significant at the 1-percent level.

The Roman numerals consist of counties grouped together to give approximately equal populations. These were theoretically equated as given in the "expected" columns. In general, the Roman-numbered groups average around 200,000 and it may be fairly easy to estimate the population of individual counties by dividing the number of counties into 200,000. Iowa has 99 counties which average around 1,000 sq. mi. in area each.

By studying the table it will be noted that discrepancies exist between expected values and observed values. The chi-square index is an evaluation of the significance of the discrepancies. A chi-square above 3.841 indicates a 5-percent level of confidence. Above 6.635 it shows a confidence level of 1 percent. Higher values show correspondingly greater probabilities of significance. It is to be noted that densely populated areas show an excess of reported accidents while sparsely settled counties show a deficiency. No speculations as to the reasons for this will be made here.

Next it seemed advisable to consider fatalities occurring within these areas since it is generally considered that accidents and fatalities are correlated, although fatalities are much more reliably reported. Table 5 shows the results obtained from this analysis.

By such technique we may set up any hypothesis that appears reasonable and check for correspondence with observed values. It might have been more logical to have used fatalities as the criterion or basis of comparison, but the results would be very similar in either case. From it we see that there is very little correspondence between the accident-reporting index and the fatality index of similar population groups with the geographical size of area varying. A further study of reasons for this observation is being made.

TABLE 5

RELATION BETWEEN REPORTED ACCIDENTS AND FATALITIES
BY AREAS USING ACCIDENTS AS THE CRITERION

Group	Number of Counties	Accidents Actual	Accidents Expected	X^2
I	1	150	344.59	109.885
II	2	138	203.40	21.028
III	3	217	232.83	1.076
IV	5	194	203.40	0.436
V	5	138	159.82	2.979
VI	8	182	136.59	13.168
VII	10	216	153.67	25.282
VIII	10	205	155.22	15.965
IX	11	175	180.09	0.144
X	12	206	127.28	48.687
XI	15	203	152.12	17.018
XII	17	160	135.04	46.135
		Total chi square		301.803

This table shows the incidence of accidents at different age levels for men. It is quite obvious that the trouble spot here lies with ages 18 to 23 inclusive. Ways and means of dealing with this group of drivers must be evolved.

TABLE 6

SIGNIFICANCE OF DIFFERENCES FOUND IN REPORTED ACCIDENTS
AT VARIOUS AGE LEVELS - MEN. EQUATED FOR MILEAGE

Age	Accidents Actual	Accidents Expected	X^2
15-17	24	16.85	3.085
18-20	103	62.85	27.324 b/
21-23	117	72.74	28.988 b/
24-26	116	99.22	3.142
27-29	103	96.32	0.498
30-32	78	87.12	1.043
33-35	52	71.64	5.789 a/
36-38	64	66.86	0.131
39-41	50	58.13	1.205
42-44	48	55.06	0.956
45-47	46	55.12	1.595
48-50	37	49.44	3.289
51-53	33	45.01	3.352
54-56	29	46.03	6.589 a/
57-59	27	37.63	3.117
60 & over	98	104.97	0.468

a/ - Significant at the 5-percent level.
b/ - Significant at the 1-percent level.

Such techniques may also be used to spot areas of poor enforcement, low accident-reporting indices, or other conditions affecting highway safety. They may be useful in diagnosing the conditions related to the incidence of accidents and fatalities. A great deal of valuable information on the causes of accidents by repeated comparisons could be gained. The method with some modification may be valuable in solving certain problems of traffic engineering.

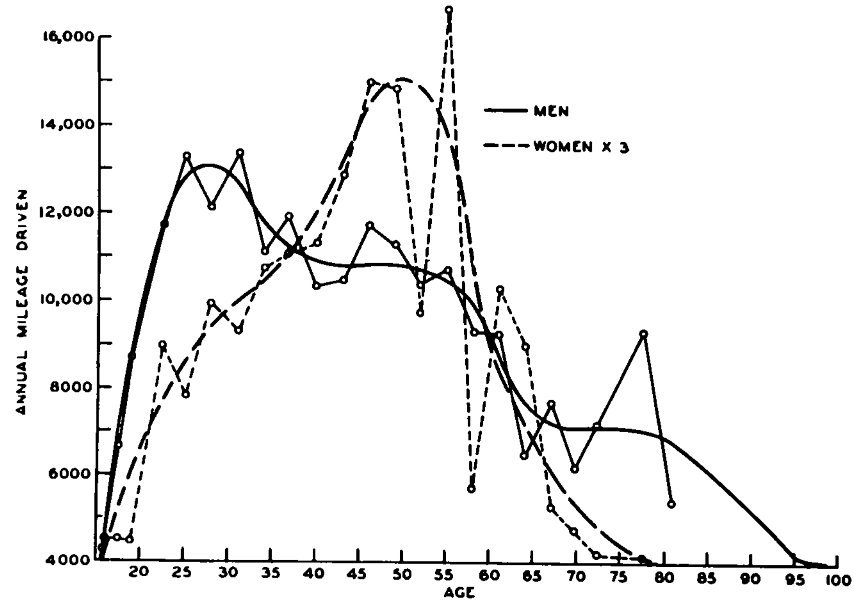

Figure 1. Relation between age and annual mileage driven (from sample of 7692 Iowa drivers; curves based on 1,419 cases replying).

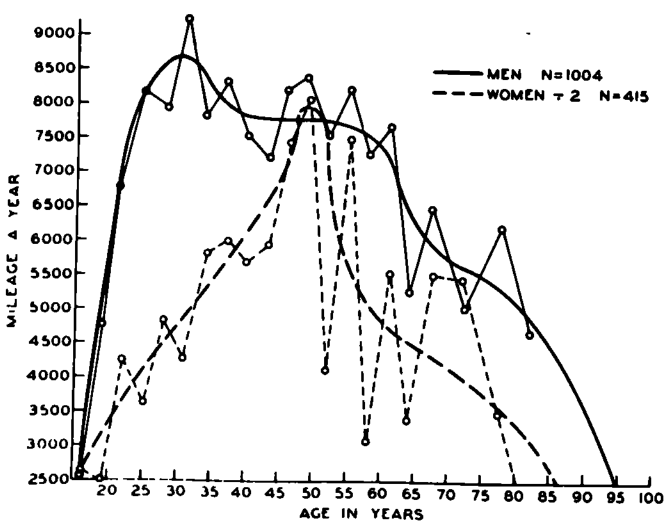

Figure 2. Relation between age and daylight driving mileages.

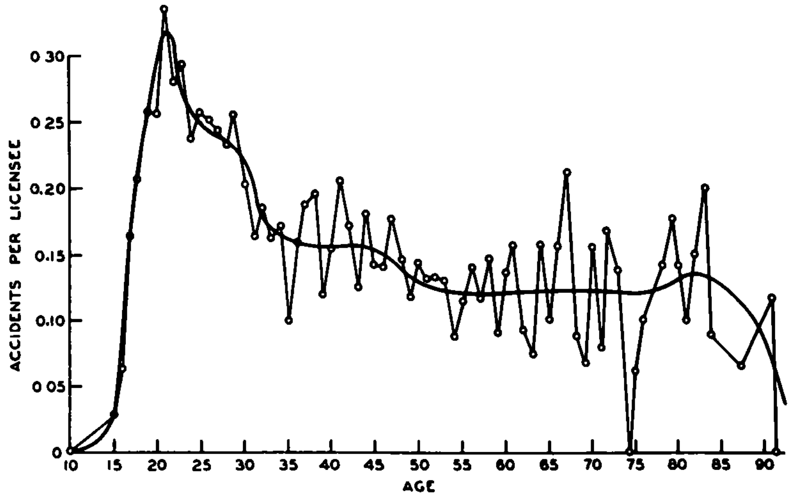

Figure 3. Accidents per licensee (1948-1949; male drivers; Iowa)

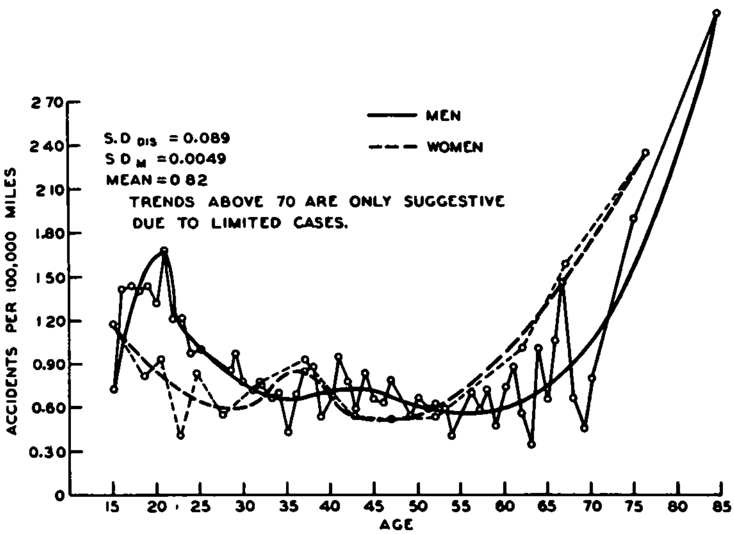

Figure 4. Accidents per 100,000 mi. (at different age levels).

These figures are more or less self-explanatory. Figures 1 and 2 show distributions of the reported mileage driven by men and women at all driving ages. The first shows annual mileage while the second indicates daylight-driving mileage. By deriving a third curve it would be possible to secure estimates of night driving at different age levels. Since a satisfactory grouping of data could not be secured for night driving, it was decided to make a separate analysis of this aspect of the youthful driving problem.

Figure 3 shows a smoothed curve of all the accidents reported for men during 1948-49 irrespective of mileage. The index is much higher for these two years than it was for any previous years as shown by Table 3.

Figure 4 shows the results obtained from plotting mileage-corrected accident indices for all age groups of men and women. It is to be noted that the index curves for men and women cross each other several times. The wider points of separation are significant. Again it must be cautioned that the mileage characteristics will likely change from year to year and such data must be kept up to date to be useful.

GENERAL SUMMARY AND CONCLUSIONS

A study of 7,692 Iowa drivers sampled from the drivers license files was made to answer two fundamental questions: Are reported accidents equally distributed among the population, age and number of licensees? Are accidents distributed equally among licensed drivers when mileage is held constant?

A number of secondary questions were posed, the answers to which may be summarized in the following conclusions subject to the nature and limitations of the study:

1. There is a preponderance of evidence that male drivers 30 and under contribute very heavily to the accident total. The differences from 18 to 23 are highly significant.

2. Male drivers spend 5 years before improvement in their reported accident record appears. Women improve their records from the beginning of their driving period.

3. Women differ from men at various age levels with respect to accidents reported against them. They drive much fewer miles a year than men and hold only about 25 percent of the licenses. They do about 10 percent of the driving and have about 9 percent of the accidents reported to the state. The chi-square test of men's and women's reported accidents was not significant, being 1.818, with a slight advantage in favor of women.

4. There is little correspondence for equal population areas, graded from most-dense to least-dense populations, between reported accidents in these areas. There tends to be an excess of accidents reported in larger cities and deficiency of reporting in sparsely settled districts, counties and areas.

5. The techniques used are suggested for spotting various conditions and situations which may be related to highway accidents.

6. Both the primary and corollary hypotheses set up for testing are rejected and the results would suggest a careful analysis of accident conditions within a state be made and used as the basis for a systematic highway accident reduction program.

7. One of the most serious problems is that of youthful male drivers. Whether the answer is probational licenses, driver education, closer surveillance and stricter enforcement, governors on cars, or other means is beyond the realm of this investigation. It appears that something will need be done, since a reasonable estimate would indicate the reduction of at least 10 percent in automobile fatalities by bringing accidents of men below 30 down to the average of other ages.

ACKNOWLEDGEMENTS

The study was made possible through a grant from the Allstate Insurance Company and set up as the Allstate Foundation for Driving Research. All possible cooperation was given by the Department of Public Safety for Iowa. The author is indebted to David E. Soles and Allen Slickers for assistance in gathering and preparing certain of the original data.

REFERENCES

Accident Facts. National Safety Council, 1946.

Fredericksen, George F. and Freeman, Herbert. "A Study of Conformity to Traffic Regulations," Journ. of Abnor. and Soc. Psychology, 1939, 34, 118-123.

Keller, H. W. "Age and Sex of Operators Licensed and in Accidents 1941-1947," Bull. State Dept. of Motor Vehicles, Hartford, Conn. (Ms.)

Lauer, Gloria L., Cutler, Max and Lauer, A. R. "Exposure Risks as a Criterion of Traffic Accident Hazards in Iowa," Proc. Iowa Academy of Science, 1945, 52, 261-263.

Moren, Mark K. and Kirklen, S. Collins. "A Study of Accident Involvement," Research Bull. Center for Safety Educ., New York Univ., 1949 (mimeo.) pp. 27.

Snedecor, George W. Statistical Methods. Iowa State College Press, 1946.

HUMAN FACTORS in HIGHWAY-TRANSPORT SAFETY

Ross A. McFarland
Associate Professor of Industrial Hygiene
Harvard University

AS OUR environment has become more mechanized and our population more mobile, accidents have increased at an alarming rate. The loss of life and incapacity from this source is as great or greater than from any known disease entity and constitutes one of the major causes of death and injury in both civilian and military populations. Existing concepts of public health, therefore should be extended to include safety at home, in the factory and in transportation from one place to another.

The control of accidents falls within the province of preventive medicine and public health because of the important role played by human variables. It is not generally appreciated that injuries, as distinguished from disease, are also amenable to the epidemiological approach and that accidents follow some of the same biological laws as do disease process. In most instances, there is multiple causation and attempts at control should involve consideration of the interaction of agent, host, and environment. Although the host is of primary medical concern, the agent and the environment must also be considered.

The extensive movement of people, goods and materials from one place to another in some form of transport vehicle is one of the outstanding features of modern life. The number of passengers using buses far exceeds those using railways, steamships or air transportation. About nine billion revenue passengers use buses each year, while somewhat less than one billion revenue passengers use the railways. Thus it appears that practically everyone uses some form of highway transport each year in the United States, not only once but many times. The American Trucking Association estimates there are approximately six million truck drivers--one of the largest job classifications in the country. The volume of supplies these drivers move from one place to another is enormous. In 1949, it amounted to about 82 million tons for intercity service by common carriers. Approximately 40 percent of the nation's freight is transported over the highways.

In spite of the many loss prevention measures which have been developed and applied in the motor transport and other industries, there are still too many accidents. Outside the home, the largest proportion of accidents occurs in some form of transport vehicle. According to the National Safety Council, each year in the United States there are approximately 40,000 motor vehicle deaths, 100,000 permanent injuries and about 1,000,000 temporary total disabilities resulting from this source. The problem of vehicular accidents is even more serious in the federal government. For each 100 million miles driven by the 360,000 motor vehicles in government service, there are 15 fatal accidents, or twice as many as the national average of 7.4. Due to the increased mileage and exposure, the actual number of deaths and injuries has increased to such an extent that the National Safety Council estimates that 10 percent of the population of the United States may be expected to be killed or injured in highway accidents in a period of about 15 years.

INVESTIGATIONS IN ACCIDENT PREVENTION

A broad research program in accident prevention is currently being conducted at the Harvard School of Public Health in the Department of Industrial Hygiene. The program is sponsored by the American Trucking Associations, Inc., the National Association of Motor Bus Operators and the National Association of Automotive Mutual Insurance Companies. It is under the direction of Dr. Ross A. McFarland, and includes in its full-time staff Mr. Alfred L. Mosely (Experimental and Clinical Psychology), Mr. Jack W. Dunlap (Psychology and Engineering), and Mr. William A. Hall (Anthropology and Engineering). Additionally, the project involves a number of part-time workers from the fields of physiology, statistics, psychiatry, anthropology, engineering, psychology, and sociology.

It is not possible here to describe all the projects undertaken by this group, but investigations are being carried out in the following areas (1) analysis of activities and distribution of work loads of bus drivers, (2) selective placement of persons by objective tests and standardized interviews, (3) health maintenance examinations, (4) study of accident repeaters, (5) human factors in the design of equipment, (6) environmental influences, (7) injuries from high decelerative forces, (8) personal adjustments and morale, (9) supervision and leadership, and (10) a study of near accidents and the critical components in driving.

DESIGNING TO SUIT THE OPERATOR

If significant progress is to be made in the future, mechanical design must be related more directly to the physiological and social habits of man. One of the first principles in accident prevention, therefore, is to design equipment to suit the operator instead of selecting an operator to fit the equipment or redesigning after the equipment is built.

All possible faults in equipment and in the working areas of vehicles should be subjected to an advance analysis in order to prevent accidents. If defects are present, it is only a matter of time before some operator "fails" and has an accident. An advance analysis involves several assumptions. The first is that an operational job analysis should include a survey of the nature of the task, the work surroundings, the locations of controls and instruments and the way an operator performs his duties. The second implies a functional concept of accidents; that is, it anticipates the errors that may occur while the operator is working at the machine. The repetition or recurrence of near or real accidents clearly indicates a need for redesign. A third consideration relates to human limitations. It should be assumed that no pilot or truck or bus driver is a perfect one. In fact, he may be far below the ability assumed by the designer. If his duties are too complex, the cumulative burden is great, and he reaches or exceeds his limits of attention and ability. Finally, a wide margin of safety should be provided to eliminate any possible situation that places the operator near his maximum ability with regard to aptitude or effort, especially when adverse factors enter the picture.

The compilation and use of anthropological measurements in relation to the design of equipment offer a very fruitful approach to the reduction of accidents. In order to improve the ease, efficiency and safety of equipment, consideration must be given to dynamic body measurements, an unobstructed view for all operations, the location of controls so that they can be worked at a bio-mechanical advantage, and restraining or protective devices in vehicles which will be decelerated rapidly. It is reasonable to demand that machines be designed from the man outward, with the instruments and controls considered as extensions of his nervous system and body appendages.

The use of "average" values to provide for the physical proportions of drivers may account for many defects. The average value cannot be employed directly, since, by definition, arrangements based upon an average would be unsuitable for 50 percent of the operators in a normally distributed group. Provision for 90 or 95 or any other predetermined percent of potential operators will require identifying the correct cut-off point. Where, for example, arm reach for the operation of manual controls is under consideration, the cut-off point should be well below the average reach; where strength of a structure intended to support one man or where body clearances are concerned, due consideration must be given to the 25 percent of the drivers whose dimensions exceed the average values as well as to the 25 percent whose measurements are below the average.

DEFECTS IN 1951 MODELS

Several examples may be given to illustrate defects in design of equipment which have direct implications for safety, comfort and efficiency:

1. Marked variations have been found to occur in the over-all working space provided for the drivers of trucks and tractors. In one instance,it was estimated only the drivers representing the smallest 40 percent of the group could be accommodated.

2. Many errors were observed in regard to human sizing. In several models, for example, only 5 percent of the drivers could comfortably reach and operate the hand brake. In others, only 60 percent could be accommodated for knee height between the pedals and the steering wheel. Some of the taller drivers are unable to adjust their sitting positions to obtain maximum visibility with regard to their instruments and the road ahead.

3. Failure to provide for adequate seat adjustments to allow for variations in human size was frequently noted.

4. There was no standardization of shift patterns in the various models studied. This predisposes the driver to error in the event that he is transferred from one model or type to another.

5. Adequate vision from the cab remains a serious problem in all models during operation in residential and busy areas. This contention is supported by an analysis of 57 light truck accidents, some 50 percent of which involved pedestrians. During bad weather, the range of forward vision is reduced approximately 50 percent, and visibility from the side is reduced even more because of no provision for cleaning or defogging.

6. Many errors have been observed in the location and design of electrical switches, especially with respect to headlamps, fog lamps and marker lights. In two instances, the dimmer switch was found to be located directly

beneath the foot pedals, behind the steering post. In these models, the driver may inadvertently operate the air horn or fog lamp while attempting to dim his headlamps. Even when he operates correctly, more complex motions and longer reaction times are required to avoid the pedals.

7. In studies of carbon monoxide concentration in trucks and tractors, the maximum safe levels were exceeded in almost every instance in which the sample was taken after the cab had been closed 15 min., with the engine idling. On repeated tests with one tractor, the level was high enough to be lethal in one hour. However, in no instance during normal cargo or passenger operations was an unsafe carbon monoxide level found. The maximum operational finding was 0.0025 percent, which is a safe level.

8. More attention should be given to the location of instruments in the panel with respect to ease of visibility. For example, in many cases the air gauge was placed directly behind the steering wheel and could not be seen by the driver without his twisting his body out of a good driving position. In another instance, the RPM indicator is placed on the extreme right of the dashboard, making it virtually impossible for the driver to read it accurately.

Design improvements alone cannot solve the accident prevention problem, and many studies point to the role of human variables in accident causation.

HUMAN FACTORS IN ACCIDENT CAUSATION

In an attempt to obtain additional information on the causes of accidents, a study was made of the records in insurance company files. The analysis included 305 cases in which a total of approximately two million dollars was paid in claims. The total number of serious accidents involved 82 drivers of fleet trucks, 63 of straight trucks, 34 of buses, and 126 long-haul drivers. The study revealed a number of significant findings. For instance, a large majority of the accidents occurred under the following conditions: at intersections or on country highways; on a dry level road surface made of concrete or macadam, having two lanes and with no reported structural defects; in daylight hours, when traffic was light. It should be noted that all of these conditions are favorable for safe operation. Human errors accounted for 90 percent of the accidents, and about 10 percent were attributable to mechanical failure. The causes which were traceable to the drivers were: (1) nonadjustment to the driving conditions, (2) inattention, (3) following too closely, (4) fatigue, (5) driving under the influence of liquor, (6) driving too fast for the range of vision and reaction time and (7) faulty meeting and passing.

PREVENTION OF ACCIDENTS BY SELECTIVE PLACEMENT OF PERSONNEL

There are several important considerations in the selection of drivers. One concerns psychological fitness, including mental ability, aptitude and temperament. Another relates to the possibility that certain employees may have repeated accidents. A third area deals with medical fitness, including such factors as vision, heart disease, susceptibility to fatigue, momentary lapses of attention or loss of consciousness, and age.

The first step in selective placement is to make an accurate job analysis. In so far as reducing accidents is concerned, this helps to identify the critical components of the job or to determine the way in which human errors may occur. An example will be drawn from the field of airline piloting. The information was obtained by "off-the-record" personal interviews as well as by direct observations. In one study it was found that 229 of 460 specific acts contributing to pilot error were due to confusion of two controls, and 83 to forgetting to operate a control. The results of these studies indicate not only that the design of equipment might be improved but also that great care must be given to the selection and training of pilots.

The next step is to develop and standardize suitable psychological tests to appraise the ability of the candidate. The tests should be given at the time of original selection to prevent labor difficulties later on. Critical scores should be determined experimentally for each job, for it is assumed that a candidate with low intelligence will be more subject to error if placed in a critical setting. It is equally important to determine cutoff points at the higher level, for it is obvious that the worker with superior ability may be bored and make errors while day-dreaming or thinking of other matters.

ANALYSIS OF THE CONCEPT OF ACCIDENT PRONENESS

The aspect of selection with which we are chiefly concerned relates to the detection of the so-called accident-prone individual. The word, "proneness" has been used to imply a series of personality traits which have not been precisely identified thus far. The term, "accident repeater" is preferable.

In any large group of workers it is often reported that a small percentage of the group give rise to most of the accident problems. In one large company, for example, it has been stated that "80 to 85 percent of the visits for injury or medical complaints are produced by 30 percent of the people. In this group there was a preponderance of neurotics and otherwise below par individuals with a high liability to hurt themselves and frequently involve others." Numerous studies from other companies including transport fields have also shown that some individuals tend to have a higher frequency of accidents than chance alone would indicate even when the exposure is controlled.

In order to demonstrate the existence of accident proneness, it is necessary to show that accidents do not distribute themselves by chance among the total group. Two lines of evidence show that this is so. In the first place, there are more people with no accidents at all and more people with multiple accidents than would be predicted by chance. Secondly, there is a tendency for those individuals who have multiple accidents in one period of time to continue to have multiple accidents in later periods of observation.

Before the importance of the purely personal factors in accident liability can be evaluated, it is first necessary to control such factors as amount of exposure in a given occupation. Furthermore, the criteria of what constitutes an accident must be controlled. If some individuals report even the slightest injuries as accidents, while others report only more serious injuries.there will be an apparent difference in liability due to this factor. Another reason which tends to make tests less successful than they might be is the multiplicity of personal factors involved. Thus, one person might have repeated accidents because of a slow reaction time, another because of a special eye defect and a third because of frequent quarrels with his wife, and so on for many other reasons.

A MAN 'DRIVES AS HE LIVES'

Our approach to the study of psychological fitness for driving relates to an exploration of the concept that a man drives as he lives. If a driver is maladjusted in his personal life, is unhappily married, is in debt, is drinking excessively and manifests other traits of personal and social maladjustments, his chances of becoming involved in an accident are much greater than if he were well adjusted. The fact that many serious trucking accidents occur toward the beginning of trips suggests that worry or apprehension carry over into the working periods, giving rise to distraction or lapses of attention. Furthermore, many drivers, who have become involved in serious accidents are known to be maladjusted in their personal lives.

In a preliminary study of high-accident as compared to low-accident taxi drivers, Dr. W. A. Tillman found that the former had a significantly high frequency of the following: divorced parents, excessive parental strictness and disharmony, excessive childhood phobias, excessive childhood aggression, truancy and disciplinary problems in school, frequent job changes, the history of being fired from employment, admitted hetero-sexual promiscuity, admitted bootlegging on the job, consciousness of physique and low interest in hobbies. Many of these items can be covered by the employment interview or identified in public records.

A study was then made of 300 passenger car operators from the general population. One hundred accident-free drivers and 100 accident repeaters were located through highway department records, and 100 additional accident-free drivers were located from insurance files. The records covered a 15-year period, and the drivers were matched for age but not for exposure. The following records were then checked: social service agency files and credit bureau, public health venereal disease records, and adult and juvenile court records. The accident repeater was known to one or more agencies in 66 percent of the cases, as compared to only 9 percent for the accident-free group.

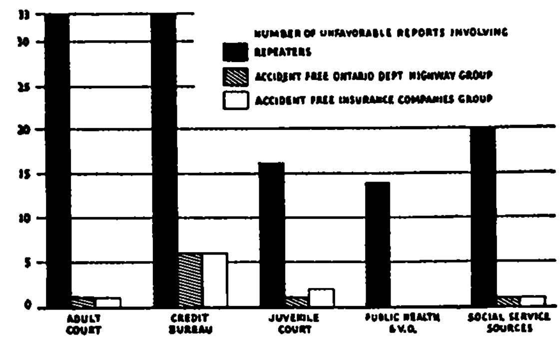

Figure 1. Social background of accident free and accident repeaters.

In our studies of truck drivers, additional refinements were made in regard to matching accident-free and accident-repeater drivers. The accident-free drivers were not involved in any accident at all, and the accident repeater had three or more chargeable accidents within a 12-month period. Exposure was controlled by matching the drivers on the following factors: (1) the traffic area in which they drive, (2) type of unit, (3) the length of the normal working day, (4) periods of operation with respect to peaks of traffic, (5) commercial driving experience, (6) marital status, (7) age, and (8) employment status. The checks made on a group of 102 drivers showed similar results to those mentioned above. The most important practical implication is that failure to check various records constitutes a serious omission in employment procedures.

PHYSICAL FITNESS

In regard to physical fitness, careful studies have shown little relationship between accidents and physical defects as revealed in the medical examinations. This may be related to the fact that the various tests are either (1) not related to the specific duties which the driver must perform, or (2) are not reliable. The physical examination for drivers should be revised on a functional basis. In other words, after the critical components of the driver's job are understood, it should then be possible to appraise those traits or functions and devise reliable tests.

One of the most important areas, which should concern the physican and safety director, relates to the basic causes of lapses of attention, distraction or temporary loss of consciousness. In many serious accidents on the highway, loss of attention may occur because of overeating, excessive fatigue or hypnotic effects of monotony. Loss of consciousness may also result from head injuries, epilepsy, advanced heart disease or diabetes. The basic causes cannot be revealed in a cursory physical examination.

If health maintenance examinations are carried out routinely and thoroughly, it is only very occasionally that a person is eliminated, and many are made safer and their useful lives prolonged if defects are detected early enough for correction.

As an illustration, consider the case of the Boston streetcar driver who was a patient of the late Dr. Soma Weiss, a specialist in the field of circulatory diseases and fainting attacks. The streetcar conductor had fainted at a particular corner on his run on several different occasions. Certain aspects of his behavior suggested that a psychological or emotional factor was involved. Dr. Weiss amazed his staff by asking to see the man's collar. He established the fact that when the conductor turned his head at the corner to look back for approaching traffic, his high collar stimulated the bundle of nerves (cartoid body) in his neck which controlled the blood supply to his brain. The man was put back to work with a low collar and had no further difficulty. This is what is meant by corrective action based on insight into the physical condition of the operator.

OPERATIONAL ASPECTS OF FATIGUE

In recent years, we have devoted a great deal of attention to the problem of fatigue in pilots, drivers and others engaged in the operation of machines. One of the most interesting findings relates to the deterioration of skills or habits which usually operate unconsciously. Studies have shown that the early signs of skill fatigue are as follows: (1) a tendency to require a stronger stimulus before action takes place, (2) a reduction in the ability to anticipate what may take place, (3) inaccurate timing of control movements, and (4) increased sensitivity, with more aggressive action both towards people and towards the machine being operated. The fringe skills or subtler reactions, which usually operate unwittingly, are lost and more effort and attention are required to carry out the tasks successfully. Finally, there seems to be a loss of insight into the extent of one's own deterioration. Such behavior naturally forms the basis for errors or accidents.

AGING AND ACCIDENTS

Not many months ago a bill was introduced into the Massachusetts State Legislature to limit the licensing of aircraft pilots to those who are 50 years of age and younger, following a fatal accident to a private pilot who was over 60 years of age. At present there is insufficient evidence to set an arbitrary retirement age for pilots or drivers. The ability to perform duties, i.e., functional age, is far more important than chronological age. Furthermore, there is a great deal of evidence to suggest that the older pilots and drivers are as safe as, or possibly safer than, younger ones. It might be possible to utilize the older worker safely if more facts were known about the functional changes which occur with age. He could then be taught how to compensate for his deficiencies. No better illustration can be given than the results which we have obtained from studies of the ability to see under low illumination in relation to age. Older people are markedly handicapped in regard to this function, and drivers over 60 or 65 years of age might be confined to day rather than night runs in order to utilize their abilities without hazard to others.

CONCLUSIONS

Our studies emphasize the fact that many variables contribute to motor transport accidents rather than any one single causative factor. Improvements can be brought about only through increased vigilance and research with regard to the design of equipment, the selection and training of personnel and safe operating practices. The prevention of accidents requires the teamwork of (1) research specialists, (2) operators who are concerned with routine scheduling and maintenance of equipment, and (3) the management groups interested in making an operating profit. It is only through coordinated efforts that any real and permanent improvement can be accomplished.

ANALYSIS of CERTAIN VARIABLES RELATED to SIGN LEGIBILITY

H. W. Case, J. L. Michael, G. E. Mount, and R. Brenner
Department of Engineering, University of California

THE RECENT development of high-speed urban freeways has increased the need for destination signs of maximum legibility. One factor thought to influence legibility is the letter-background arrangement, with the usual alternatives being black letters on a white background and white letters on a black background. A survey of the literature revealed nine published studies in which the legibilities of these two combinations were systematically compared.[1] The results of these studies throw little light on the problem as it relates to highway destination signs.

All but one of the studies were conducted in the laboratory with illumination levels ranging from 1 to 100 foot-candles. Since outdoor daylight may range as high as 10,000 foot-candles, results obtained in lower illumination levels may not be directly applicable to the usual daylight situation. However, if the results are considered only in terms of the low levels of illumination, the problem still remains unresolved because of the inconclusive and contradictory results obtained. Furthermore, the earlier experiments were, in general, based on the principle that all variables but one should be held constant. Therefore, it was assumed that the relationship between letter-background arrangement and legibility was independent of other factors. The failure to assume interaction between several relevant variables may be responsible in part for the contradictory nature of the findings.

As a result of the survey and analysis of the experimental literature related to the subject, it was decided that an investigation should be made in a field situation to attempt to determine whether or not outdoor, daylight letter legibility varies with letter-background arrangement, and in turn how this letter-background legibility is influenced by spacing between letters and rows.

CONSTRUCTION OF THE TEST SIGNS

Letter-background Arrangement. Using the Bureau of Public Roads Standard Alphabet for Highway Signs, 1945, as a model, 3-in.-high Series E (stroke width approximately one sixth of letter height) rounded capital letters were carefully drawn and then contact photographed. Contact prints were made from the negative and from a paper positive on white, semimat (G-81 contact) photographic paper.

The letters B, C, D, E, F, O, P, and R were used for the experiment. These particular letters, which appear in the American Optical Company Snellen Chart for visual acuity, were considered appropriate for this experiment.

1/ - Details of these nine studies are presented in the annotated bibliography starting on page 55.

From the black and white letter reproductions, 40 random groups of eight letters were then chosen so that for each "randomly" selected group of eight black letters there was an identical group of eight white letters. However, the order of appearance of the letters was not the same for both the black and the white card of a particular pair. In addition, all of the eight different letters used did not necessarily appear on any single card, since the selection method resulted in some letters appearing more than once in a specific group. Thus, there were 20 matched pairs of letter groups.

Spacing. The matched pairs of letter groups were then fastened on black or white mat boards, 20 by 30 in. Eight letters appeared on each board in two rows of four letters each. Ten of the 20 paired groups were placed on the boards so that there was 1½ in. (half the letter height) between each of the letters in a row of four, and 2 in. between the two rows. The other ten pairs were spaced 4 in. between letters (1 1/3 of the letter height) and 4 in. between the two rows.

The experimental material, therefore, consisted of 20 black signs and 20 white signs. Half of the black and half of the white signs contained eight closely[2] spaced letters; the other halves contained eight widely spaced letters. This gave a matching of pairs in which each member of the pair had the same spacing.

PLACEMENT OF THE SIGNS

Several problems arose in determining the most adequate distance of the signs from the subject for best observation purposes. Undoubtedly the viewing range is important, for if the signs are placed at distances such that the subject can read either none of the letters or all of them, no legibility differences resulting from the black and the white letters can be expected. It is also not known whether the effect being studied remains constant over the range between these extremes, or manifests itself differentially for different degrees of difficulty, i.e., different distances.

The problem of cooperation on the part of the subject became an important factor since the experiment involved a restricted routine task taking approximately ½ hr. of the subject's time. If cooperation was to continue throughout the observation period, the task should permit a reasonable number of successes.

In a preliminary study an attempt was made to find a single range composed of five equally spaced distances, so that none of the subjects would read less than half of the letters correctly and none would get more than 80 percent correct. A single range containing observation distances meeting these requirements was not found. However, it was possible to establish two separate ranges. When the subjects with better visual acuity were tested on the distant range and the subjects with poorer visual acuity were tested on the close range, each range was found to be satisfactory for its appropriate group of subjects. The distant range consisted of five equally spaced positions between 295 and 355 ft. The close range consisted of five equally spaced positions between 220 and 280 ft.

2/ - Actually, these closely spaced letters were as far apart for their height as those on any highway destination sign.

SIGN PRESENTATION

Each range contained five distances, and since there were 40 signs in all, eight signs, or four matched pairs, were shown at each distance. The letters on two of the matched pairs were closely spaced and those on the other two were widely spaced. Thus, any differences between white letters and black letters at a particular distance would not be due to differences in letter difficulty.

To exclude the possibility of presentation order affecting the results, a different random order of the 20 pairs of signs was used for each subject. Both members of a matched pair appeared at the same distance; and at each of the five distances, half of the signs were closely spaced and half were widely spaced. The order of appearance of the five distances was randomized for each subject.

SELECTION OF SUBJECTS

Subjects were initially screened for visual acuity of 20/40 or better on the Keystone Industrial Telebinocular No. 46A. For further refinement of measurement of the visual acuity levels of each subject a large-scale eye chart was constructed using the same letters (both black and white) as those used in the experiment. This was scaled on the basis of the results of preliminary measurements. There were 28 letters on this chart. If a prospective subject could read 19 or more of the letters at a distance of 280 ft., he was assigned the more distant range in the experiment; if he read 11 to 19 letters, he was assigned the closer distance; and if he was unable to read 11 letters, he was not used in the experiment.

All of the subjects were volunteers from enrolled students at the University of California, Los Angeles, Summer Sessions of 1951.

FIELD LAYOUT

A large, flat area which had just been leveled for construction purposes was chosen as the experimental site. An apparatus consisting of an exposure shutter was placed at one end of the field and a movable sign holder at the other. The two ranges on which the signs were to be shown were marked off on the field.

The shutter-like apparatus was painted olive drab, and the background against which the signs were shown was the light brown of recently turned earth.

EXPERIMENTAL PROCEDURE

One experimenter and one subject sat behind the shutter device. Another experimenter stood beside the sign holder, which had been moved to the first position at which this subject was to see the signs. A decision was made as to whether the close or distant range was to be used on the basis of the prior acuity measurements.

A number of small cards had been prepared showing random sign orders and distance orders. Before the subject's arrival, the 40 signs had been stacked according to one of these cards. All 40 signs were placed in the sign holder so that when one was removed the next one in order was revealed.

The subject was told what letters were being used (B, C, D, E, F, O, P, R) and was handed a small card on which these letters were shown.3/ The shutter was then opened and the subject instructed to read the letters at a rate which suited him, and to say "blank" when a letter was completely unrecognizable. The experimenter at the shutter recorded the letters as the subject read them. When the subject finished with a sign, the experimenter signalled the man at the sign holder to remove the observed sign, thus revealing the next one.

When all eight signs at a particular distance had been read, the shutter was closed, the sign holder moved to a new position, and the shutter reopened. This procedure was continued until the subject had seen eight cards at each of five positions.

No instructions were given as to rest periods or means of reducing eye strain. The subjects were free to view the signs in whatever manner they chose. Nor were they told the exact nature of the experiment; however, this was obvious to most of them after a few of the signs had been shown.

STATISTICAL TREATMENT OF RESULTS

The results consist of four scores for each subject. These scores are the number of letters correctly identified for each of the four different kinds of signs. Since there were ten signs of each kind and eight letters on a sign, the maximum possible number of correct letters is 80.4/ Thus, in Table 1, for subject 1, the score 36 means that Subject 1 correctly identified 36 of the 80 closely spaced white letters on a black background.

With two experimental variables and two degrees or qualities of each, a two-by-two factorial analysis of variance with replication was selected as an appropriate analytic design. Since each subject experienced all four experimental conditions, the number of replications is equal to the number of subjects.

From the subject totals (Column 6, Table 1) it is apparent that a fairly wide range of visual acuity was sampled. However, if the subjects are grouped with respect to visual acuity, it is possible to examine one more factor, since it is not known whether the relationships between letter-

3/ - This was due to finding, through preliminary measurements, that the subjects tended to make consistent errors due to erroneous first judgments. For example, if it happened that the subject's first sign was seen at such a distance that he misread the R, calling it an N, he would continue to make this error even at closer distances. It appeared that this was due to careless reading, since if urged to study the letter carefully at the closer distances, the subject would correct the error.

4/ - Each of these four scores actually represents sums of correct letters for all five distances even though distance is orthogonal to letter-background arrangement and spacing. Distance was not analyzed as a separate factor, because the wide individual differences in visual acuity render it relatively meaningless.

TABLE 1

Subject	Close Spacing white letters	Close Spacing black letters	Wide Spacing white letters	Wide Spacing black letters	Subject's Total for all letters
1	36	43	61	58	198
2	26	41	49	56	172
3	28	33	51	42	154
4	38	43	66	66	213
5	42	52	63	67	224
6	34	36	55	54	179
7	56	68	72	69	265
8	36	32	51	64	183
9	37	47	61	63	208
10	19	22	60	47	148
11	28	45	55	61	189
12	45	47	70	68	230
13	20	29	35	34	118
14	25	28	50	38	141
15	27	24	52	45	148
16	31	30	59	46	166
17	55	56	64	68	243
18	23	40	46	50	159
19	57	57	70	71	255
20	40	43	57	57	197
21	49	52	62	59	222
22	40	51	60	61	212
23	32	36	57	63	188
24	40	49	58	63	210
25	47	40	54	54	195
26	34	38	55	53	180
27	54	39	72	67	232
28	22	26	46	44	138
sums	1021	1147	1611	1588	
means	36.5	41.0	57.5	56.7	

Combined Means

white letters on black background	94.0
black letters on white background	97.7
close spacing	77.4
wide spacing	114.2

background arrangement, spacing, and letter legibility are independent of visual acuity (or reciprocally, independent of the difficulty of the task). The grouping was accomplished by ranking the subjects with respect to total correct letters and dividing the 28 subjects into four groups of seven, as shown in Table 2. Thus, the analysis of variance design is a two-by-two-by-four analysis with replication.

The calculations for the two-by-two-by-four analysis were accomplished in the usual manner[5] with the results shown in Table 3.

One of the assumptions of analysis of variance is that the variables being dealt with are distributed normally. A chi-square goodness-of-fit test was made and the distribution was found not to differ significantly from a normal distribution with the same mean and variance.

Another assumption made when using analysis of variance to test the significance of the differences between means is that the variances are homogeneous. This assumption was tested using Bartlett's method[6] and found to hold. The uncorrected value for chi squared with three degrees of freedom was 2.21. The probability of obtaining this value, or a larger one, is greater than .50.

INTERPRETATION OF RESULTS

Spacing. In Table 3 it is seen that a highly significant *F* was obtained for the difference between the close and wide spacings. The means shown in Table 1 reveal that the widely spaced letters were much more legible than those closely spaced. A possible explanation in terms of fixation shifts is as follows.

Due to head movements, eye movements, and instability in the visual mechanism, the image of an object shifts about slightly on the retina, even though the observer may try to keep his head and eyes still. When there is wide spacing between two objects, the only effect of this shifting is a slight blurring of the edges of the objects. If the spacing between them becomes sufficiently small, the net result of the shifting is a superposition of adjacent letter parts which could decrease letter legibility.

Interaction Between Letter-Background Arrangement and Spacing. The highly significant *F* obtained for the interaction of the two variables means that the legibility of the two letter-background arrangements is not independent of the letter spacing. Black letters on white background is the most legible arrangement when the letters are closely spaced, but this is not true when the letters are widely spaced. It may even be that white letters are superior when wide spacing is involved. A possible explanation may be found in the phenomena of irradiation and the fixation fluctuation mentioned above.

5/ - Edwards, Allen L., Experimental Design in Psychological Research, Rinehart and Co., Inc., 1950, pp. 237-260.

6/ - Eisenhart, C; Hastay, M. W.; Wallis, W. A., Techniques of Statistical Analysis, McGraw-Hill Book Co., 1947, pp. 387-388.

TABLE 2

Visual Acuity	Close Spacing: White Letters	Close Spacing: Black Letters	Wide Spacing: White Letters	Wide Spacing: Black Letters	Subject's Total for all letters
Group I	56	68	72	69	265
	57	57	70	71	255
	55	56	64	68	243
	54	39	72	67	232
	45	47	70	68	230
	42	52	63	67	224
	49	52	62	59	222
Group II	38	43	66	66	213
	40	51	60	61	212
	40	49	58	63	210
	37	47	61	63	208
	36	43	61	58	198
	40	43	57	57	197
	47	40	54	54	195
Group III	28	45	55	61	189
	32	36	57	63	188
	36	32	51	64	183
	34	38	55	53	180
	34	36	55	54	179
	26	41	49	56	172
	31	30	59	46	166
Group IV	23	40	46	50	159
	28	33	51	42	154
	19	22	60	47	148
	27	24	52	45	148
	25	28	50	38	141
	22	26	46	44	138
	20	29	35	34	118

TABLE 3

ANALYSIS OF VARIANCE

Source of Variation	df	S S	M S	F	P
within	96	2617.58	27.27		
black and white	1	94.73	94.73	3.47	grearer than .05
spacing	1	9490.72	9490.72	348.03	less than .001
visual acuity	3	8451.53	2817.18	103.31	less than .001
black and white by spacing	1	198.22	198.22	7.27	less than .01
black and white by visual acuity	3	74.66	24.89	. . .	
spacing by visual acuity	3	147.53	49.18	1.80	greater than .10
black and white by spacing by visual acuity	3	84.03	28.01	. . .	
Total	111	21159.00			

Column Heading Legend

df = degrees of freedom.

S S = sum of squared deviations from the mean.

M S = mean sum of squared deviations.

F = F ratio

P = probability of obtaining an F as great or greater.

A bright object on a dark background appears to be wider than a similarly sized dark object on a bright background. Though the underlying casual factors are not agreed upon7/, this phenomenon has long been known and is called irradiation. Such irradiation was easily observed in the case of our letters. The white ones appeared to be considerably wider than the black ones. Conversely it may be that the black letters were affected in an opposite manner, and seemed narrower due to the irradiation of the white areas surrounding them.

The overall effect of this irradiation is an increase in the apparent stroke width of the white letters and therefore a decrease in the apparent spacing between letters. The stroke width of the black letters appears to either remain constant or actually decrease, thereby increasing the apparent spacing between the letters. Thus, even with constant distances between letters, the apparent spacing is decreased for the white letters and possibly increased for the black letters. It follows that the white letters would be expected to suffer more interference due to fixation shifts and therefore to be less legible.

As spacing increases in magnitude, it eventually becomes so great that interletter interference does not occur, or at least is not the limiting factor affecting the legibility. For widely spaced letters then, an increase in the apparent stroke width of the white letters would not necessarily decrease legibility. In fact, if the letters were viewed at such a distance that the stroke widths became the limiting factor, the white letters might be expected to be more legible than the black letters, due to their greater apparent stroke width.

It will be remembered that letter-stroke width remained constant during the experiment. Therefore, the results, strictly speaking, are limited to this stroke width. However, if this explanation of the results is valid, it would seem quite possible to predict what the effects of varying stroke width would be.

For the closely spaced letters we would expect an increase in legibility with a decrease in stroke width for both black and white letters. This would be due to effectively increasing spacing and, therefore, decreasing the interference between the letters which is caused by fixation fluctuations. Eventually a point would be reached where further decrease in stroke width would decrease legibility. It is at this point that stroke width itself becomes the limiting factor in legibility. However, we would expect this maximally legible stroke width to be smaller for white letters than for black letters, due to irradiation.

With the closely spaced letters, fixation fluctuations and irradiation are thought to be of considerable importance, as these may be the cause of interference between the letters. In the case of the widely spaced letters, however, consideration must be given to the overall shape, size of crucial distances between letter parts, knowledge of what letters are being presented, and other factors somewhat more complex. The manner in which these factors interact with irradiation and fixation fluctuation will depend on

7/ - Bartley, S. Howard, Vision, D. Van Nostrand Co., Inc., 1941, p. 235 ff.

many additional factors including the design of the letter itself. For these reasons, any general inferences with respect to the widely spaced letters on the basis of this one experiment would be unwarranted.

Visual Acuity or Difficulty of the Visual Task. It is interesting to note that there were no significant interactions involving the variable of visual acuity. Hence, the other relationships investigated in this experiment appear to be reasonably independent of visual acuity or the difficulty of the task for the general conditions of this experiment.

Letter-Background Arrangement as a Main Effect. A sizable F was obtained for letter-background arrangement as a main effect. The means in Table 1 indicate that black letters on white were better than the opposite arrangement. If the spacings used had been selected randomly, this effect might be interpreted to mean that, in general, black letters on white are superior to white letters on black. Actually, no such interpretation should be made, since the spacings were not chosen at random and the significance of this effect is due entirely to the result of close spacing.

SUMMARY

Black letters on a white background and white letters on a black background were compared with respect to legibility. Both closely spaced and widely spaced signs were used. The subjects were divided into four levels of visual acuity, on the basis of the results. The results indicated that the widely spaced letters were more legible than those closely spaced, and that the legibility of the black-white arrangement as compared with the white-black arrangement was dependent on the letter spacing. This interaction was discussed in terms of irradiation and fixation fluctuations, and these concepts were then used to infer the effects of decreasing stroke width. These inferences were limited to the closely spaced letters.

REFERENCES

Aldrich, Milton H., "Perception and Visibility of Automobile License Plates", Proceedings, Highway Research Board, Vol. 17, pp. 393-411, (1937).

Forbes, T. W., "A Method for Analysis of the Effectiveness of Highway Signs", Journal of Applied Psychology, Vol. 23, pp. 669-684, (1939).

Forbes, T. W., and Holmes, Robert S., "Legibility Distances of Highway Destination Signs in Relation to Letter Height, Letter Width, and Reflectorization", Proceedings, Highway Research Board, Vol. 19, pp. 321-335, (1939).

Forbes, T. W.; Moskowitz, Karl; and Morgan, Glen; "A Comparison of Lower Case and Capital Letters for Highway Signs", Proceedings, Highway Research Board, (in press).

Janda, H. F., and Volk, Wayne N., "Effectiveness of Various Highway Signs", Proceedings, Highway Research Board, Vol. 14, pp. 442-447, (1934).

Lauer, Alvah R., "Improvement in Highway Safety", Proceedings, Highway Research Board, Vol. 12, Part I, pp. 389-399, (1932).

Lauer, A. R., "The Effect of Increased Illumination on Acuity Readings of Singly Exposed Letters at Different Distances", Year Book of Optometry, 1942, pp. 225-232.

Otuka, J., and Honda, T., "Relationship Between Visual Acuity and Contrast Under Several Light Intensities", Acta Societatis Ophthalmologicae Japonicae, Vol. 44, pp. 2253-2259, (1940).

Wilcox, Warren W., "The Basis of the Dependency of Visual Acuity on Illumination", Proceedings, National Academy of Sciences, Vol. 18, pp. 47-56, (1932).

ANNOTATED BIBLIOGRAPHY

Berger, Curt, "Stroke-Width, Form and Horizontal Spacing of Numerals as Determinants of the Threshold of Recognition," Journal of Applied Psychology, Vol. 28, pp. 208-231, (1944).

Threshold distances were found (for single numbers and number groups) by moving a sign toward the observer until he could recognize it, and moving it away until he could no longer recognize it. The outer area of numbers was kept constant (80 mm. by 42 mm.) while the stroke width was varied. Three numbers (8, 5, and 2) were used in a comparison of black and white legibility. The optimal stroke width for white was found to be 6 mm., and that for black was 10 mm. When these optimal widths were compared, it was found that the white was better by about 9 percent. Two subjects of normal vision were used. No measures of variability or reliability are presented.

Holmes, Grace, "The Relative Legibility of Black Print and White Print," Journal of Applied Psychology, Vol. 15, pp. 248-251, (1931).

The relative legibility of white and black print are studied by the "distance method." Isolated 5-letter words (64 words in all) were mounted, 16 at a time, on $8\frac{1}{2}$-by-11-in. cards. Four of the cards were uniformly lighted with an intensity of 35 foot-candles per square foot. Each card was started at one end of a bench with the subject seated at the other end. He was instructed to read as many of the words as he could. The card was then moved forward 20 centimeters. The procedure was continued until the subject had correctly read all 16 words on each card. Twenty subjects were used. The distance at which each word was correctly read was recorded and the means of these distances for the white and black print were compared. The difference was found to be highly significant. All relevant variability and reliability data are given.

Kirschmann, A., "Uber die Erkennbarkheit Geometrischer Figuren und Schriftzeichen im Indirecten Sehen," Achiv fur die Gesante Psychologie, Vol. 13, pp. 352-388, (1908).

Utilizing a campimeter he had geometric figures and letters presented to one subject. The criterion of legibility was the distance in centimeters from the fixation point at which the letter or figure could be correctly identified. For the white stimulus object the total background was black and vice versa for the black stimulus object. The illumination was 5 foot-candles throughout the experiment. Each alphabet was presented several times. Averaging the distances for all letters and figures, it was found the white on black could be seen 16 percent further from the fixation point than the black on white.

Lauer, A. R., "Certain Structural Components of Letters for Improving the Efficiency of the Stop Sign," Proceedings, Highway Research Board, Vol. 27, pp. 360-371, (1947).

After a discussion and evaluation of past results dealing with sign effectiveness, the author presents a description of 7

series of experiments carried out at Iowa State College. These were titled: (1) Differential Perceptibility of Certain Letters, (2) The Optimal Spacing of Block Letters, (3) Optimal Background Letter Combinations, (4) Optimal Letters' Shapes, (5) Relative Perceptibility of Sign Shape and Color, (6) Comparison of Thresholds for Visability, Color, and Shape, (7) Comparison of Laboratory and Outdoor Results on Typical Signs Used. The method used to compare black and white letters was one of gradually increasing the illumination (in the laboratory situation) until the letter was perceived correctly. The comparison was in terms of light units required for accurate perception. The results indicated that black letters required about 20 to 25 percent less light units than white letters. This method was also used in series 2. The author states that "in all regular experiments the conclusions given are based only on cases shown to be significantly different." However, no measures of significance are presented and the level of significance considered is not reported.

Mills, F. W., "The Comparative Visibility of Standard Luminous and Non-luminous Highway Signs," Public Roads, Vol. 14, No. 7, pp. 109-128, (1933).

The workers at the Bureau of Standards attempted to investigate the visibility of standard traffic signs (STOP, SLOW, CURVE, etc.). A tachistoscopic arrangement was used to limit the time of observation. Five signs (each of different dimensions) were used at each of 5 distances (200 to 500 ft.). The signs were of three different shapes. Three sets of the five signs were made: black letters on white, black letters on yellow, and white letters on black. These three were compared in terms of percentage of correct identifications for each distance. Between 10 and 20 subjects were used at each distance and a total of from 600 to 800 observations were made at each distance. With respect to the black on white and the white on black, the results were inconclusive. It is difficult to tell from the report exactly what procedure was followed in the exposure of the signs. No variability measures of any kind are presented. The study also dealt with night observations on non-luminous signs, night and day observations on luminous signs (three different sizes and shapes of reflector buttons), and recognition of a sign from its shape.

Paterson, Donald G., and Tinker, Miles A., "Studies of Typographical Factors Influencing Speed of Reading: IV Black Type vs. White Type," Journal of Applied Psychology, Vol. 15, pp. 241-247, (1931).

Black print on white and white print on black are compared with respect to speed of reading. A control group read two speed-of-reading tests, both printed black on white, while an experimental group read the same tests with the second printed white on black. Each group contained 140 subjects. A small significant difference in favor of the first test was found for the control group, whereas, a large significant difference in favor of the first test (black on white) was found for the experimental group. On the basis of these figures it is estimated that there is about a 10.5 percent difference in legibility as measured by reading speed in favor of the black print on white ground. The standard formula for determining the statistical significance of obtained differences for correlated measures are employed. All relevant variability and reliability data are given in tabular form.

Scott, Walter D., The Theory of Advertising, Small Maynard and Co., 1907, pp. 138-140.
Specially prepared cards containing combinations of white letters on black and vice versa were shown individually to a number of subjects for one-seventh of a second per card. The results showed that black letters on white were seen oftener than the white type on black.

Starch, D., Advertising, New York, Scott, Forsman, and Co., 1914, pp. 189-190.
Forty persons were given two tests of reading speed. One test was black on white and the other was white on dark gray. The average number of words per second was 4.26 for the white on gray and 6.06 for black on white.

Taylor, Cornelia D., "The Relative Legibility of Black and White Print," Journal of Educational Psychology, Vol. 25, pp. 561-578, (1934).
Four different measures of legibility were used. These were (1) the span of visual apprehension method of Scott, (2) Kirschmann's legibility in peripheral vision, (3) a photographic record of eye movements during reading, and (4) the distance method of Holmes. Four separate experiments are involved, each requiring separate description. In the first experiment, 3-by-4½-in. block capital letters were cut out of white, light gray, dark gray, and black paper and pasted on 7-by-28-in. cardboards, nine letters at a time. The cardboard was black for the white letters and white for the other three colors. These cardboards were then exposed for three seconds each and the number of letters correctly identified was recorded. Each brightness combination appeared four times. The apparatus was a projection type apparatus making it possible to present the material simultaneously to 30 subjects at a time. The total N was 128. Mean correct identifications (out of a possible 36 for each brightness combination) were as follows: black on white, 24.63; dark gray on white, 24.70; light gray on white, 23.29; white on black, 23.93. The standard errors of these means were all about 0.31. The differences between means are all significant above the 0.05 level except the black, dark gray comparison. Standard errors of correlation differences were used and the correlations were given.

The second experiment was an attempt to duplicate Kirschmann's experiment. Three slightly different white alphabets were compared with the black one. Six subjects read each white alphabet in comparison with the black alphabet. The presentations were distributed over a three day period and mixed so as to control practice and fatigue effects. The mean distance from the fixation point for all three white alphabets was 11.66 cm. The mean distance for the three readings of the black alphabet was 9.32 cm. This difference between these means is significant above the 0.01 level and in the opposite direction from the results of Kirschmann.

In a third experiment, Taylor utilized photographs of eye movements to compare reading performance on white-and-black print. Ten subjects were used and each read 2 sections from form A of a reading test (black on white) and 2 sections from form B (white on black). All four sections were read at a single sitting but they were rotated to control fatigue and practice effects. Three of the four measures of reading performance showed significant differences (S.E. of differences between correlated means) in favor of the black print while one showed an insignificant difference in favor of the white print.

The last experiment is really a series of several experiments all utilizing the method and apparatus of Holmes with a constant illumination of 100 ft. candles. Ten subjects were presented with six different sets of stimulus cards. The factors varied were size of print, type of print (Scotch-Roman, simple block, and Kabel lite), and meaningfulness of the material. For type size, black on white was superior for all sizes. With respect to type face, black on white was significantly superior for Scotch-Roman face, whereas for Kabel the results were insignificant. The superiority of black on white was found to increase as the meaningfulness of the material decreased.

35-MILLIMETER AIRPHOTOS for the STUDY of DRIVER BEHAVIOR

T. W. Forbes* and Robert J. Reiss**
Institute of Transportation and Traffic Engineering
University of California, Los Angeles

SYNOPSIS

IN 1940 an attempt was made by one of the authors to use 16-mm. movies from a blimp to study driver behavior on the highway. Results were not satisfactory. This report deals with another attempt to obtain a rapid sequence of airphotos cheaply enough to make their use feasible, using a small, compact, relatively cheap, 35-mm. movie camera and a light airplane. The camera was slightly modified, and a single-frame electric drive was added so that it could be mounted externally, if desired, thus obviating the need of a special plane and allowing use of an electric interval timer. Automatic recording of a sweep-second watch and counter was provided on the film.

For least error of parallax, photos should be taken as nearly vertically as possible. In order to follow the turns in a highway, however, it is necessary to bank the airplane. To make this possible with the device mounted externally, a self-orienting-and-aiming mechanism was planned, but the project was terminated before it was built. The camera was tried out in the air by means of a metal frame and cowling, which allowed it to be placed on the step of a light plane flown with the door off at about 4,000 feet of altitude. Pictures were obtained which were sharp enough to project, thereby easily enlarging the image by as much as 40 diameters to show traffic clearly.

Although further work is necessary on the interval timer, aiming mechanism, and procedure for keeping vertically over the highway, the method has been shown to be practical. Such pictures can obtain, in a short time, information on the adequacy of interchange design from driver-vehicle behavior and on passing practices and other driving procedures. The method should also lead to the development of longitudinal or "sweep" sampling of street and highway traffic as contrasted to the single location or "ambush type" of sampling.

IN 1928 use of aerial photographs for traffic studies was made in the Maryland Aerial Traffic Survey (1). A mapping camera was used to obtain aerial photos for a traffic volume count on the Baltimore-Washington Highway.

---

*Now executive secretary of the Committee on Highway Safety Research of the National Research Council, National Academy of Sciences, Washington. All interpretations and conclusions are those of the first author personally and are not necessarily those of any other persons or organizations involved.

**Now shop superintendent, Enterprise Iron Works, Los Angeles, who designed and built the electric drive and the time-and-frame-count recording chamber.

In 1939 Forbes (2) described use of air photos made with a 16-mm. movie camera for the study of driver behavior on the highway. Subsequently a flight was made in a Goodyear blimp from New York to New Haven over the Merritt Parkway, in which pictures were taken at 8 frames per second with a 16-mm. Kodak Cine Special. Results were not satisfactory, however, and were not reported.

In 1947 Greenshields (3) reported an attempt by Shapiro and himself to use a movie camera from a helicopter and, later, from military aircraft. Both were unsatisfactory, and it was concluded that the use of a standard mapping camera with a larger size of film was necessary.

In 1948 two helicopter flights were made for the Dayton Plan Board (4) using a 16-mm. movie camera for obtaining parking-survey data and traffic-volume counts. This attempt was not entirely satisfactory but was said to show that the method was feasible for such surveys.

The use of airphotos for mapping, highway location, and similar applications has become an established field and has been treated by numerous authors. Mapping cameras have also been used for traffic studies since 1928. A K24 aerial-mapping camera, which utilizes 5½-in. film was used by Terry (5) from a light plane in 1949, and various others have used cameras taking large-size pictures, both hand-held and rigidly mounted.

SHORTCOMINGS OF PREVIOUS METHODS

Although satisfactory for mapping, alignment, and some traffic studies, the standard methods of aerial photography are unsatisfactory for studying the behavior of drivers on highways. The use of mapping cameras requires large-sized, expensive film. Furthermore, most mapping cameras are limited to a maximum of 120 to 150 pictures on a roll. A longer sequence would require rapid change of several magazines, so a long series of pictures taken at short intervals would not be very practical.

The necessity for the use of a special airplane with a camera port in the floor limits availability of equipment to many organizations, and if a specially equipped plane must be brought from one location to another to take the pictures, it is often impossible to take advantage of the good weather necessary for the pictures.

REQUIREMENTS FOR DRIVER BEHAVIOR STUDIES

In order to follow the behavior of drivers on the highway, a long series of single exposures is required. These should be spaced at a short interval (down to 1.0 sec.). For least parallax the pictures should be cheap enough to be within the reach of the universities, research laboratories, and similar organizations which might carry on traffic engineering and highway safety research. And to be of maximum usefulness, the method should be usable in standard light planes available in any part of the country.

SELECTION AND MODIFICATION OF A CAMERA

Accordingly, a standard 35-mm. movie camera of a less expensive make was modified for aerial use. The camera unit was kept light enough so that it could be used with any high-wing, four-place light plane. Such a plane can be rented for from \$8 to \$15 an hour with another \$3 or \$4 if the pilot is also hired. Such planes are available at almost any small airport in the United States, and a skillful pilot can maneuver such an airplane over a highway in slow flight without too much difficulty (except possibly over crooked mountain roads).

A De Vry 35-mm. Model A news camera was chosen. This camera is ordinarily spring driven and used by hand for news photos where a larger, more-complex camera cannot be used. It uses a 100-ft. roll of 35-mm. movie film, which provides 1,200 individual exposures.

This camera is equipped with an extra shaft which allows hand operation when it is disconnected from the spring-drive mechanism. An electric-motor drive with a single-frame exposure mechanism was built to drive this shaft. Thus a drive was provided which could be remotely controlled and which, therefore, would allow mounting the camera outside the plane if desired. An electrical interval timer for spacing the photos could also be used. An exposure counter and clock were built into the camera unit in such a way that the frame count and time were thrown onto the film automatically. By a combination modification of shutter and drive speed, the camera was given exposure speeds suitable for aerial photography.

For the electric drive, a 24- to 28-volt D.C. motor from a surplus aerial gun camera was utilized. These motors are high-speed, governor-controlled motors with a step-down gearing unit built into them. Further step down was obtained by gearing.

Fast operation of the camera shutter required that the motor be continuously operated, since otherwise it would take a considerable interval to bring the camera up to speed. In order to obtain such operation the motor was run continuously and was coupled to the drive shaft of the camera through a planetary gearing mechanism. The motor drive shaft was connected to the sun gear, while the planet gears were mounted on a plate which was attached to the drive shaft. A braking mechanism operated on the ring gear of the planetary drive. A solenoid operated this brake to start the camera (see Figs. 1 and 2).

A hook-and-detent mechanism was devised in order to make single-frame exposures positive when a single, short impulse was sent to the solenoid. A second solenoid raised this hook just previous to the operation of the brake solenoid. The brake solenoid then started the rotation of the camera drive shaft. The latter was stopped by the hook which dropped into a detent when the trip solenoid was released and an exposure was completed.

AUTOMATIC RECORDING OF EXPOSURE NUMBER AND TIME

The De Vry camera is fitted with a small, movable plug to allow viewing a scene through the lens if desired. Such viewing is made possible by a small mirror mounted directly behind the film gate. Advantage of this was taken in order to throw the image of a watch and a three-digit counter onto the back of the film at one corner of the exposure. This area was then masked from further exposure by a slight addition to the film gate in front of the film. A standard watch with black face and white hands, such as is used in aerial cameras, and a three-digit, 24-volt counter were mounted in a chamber which was fastened to the outside of the camera. This chamber was light proof and was fitted with a mirror and lens in such a way as to focus an image of the watch and counter onto the mirror inside the camera and thence onto the film.

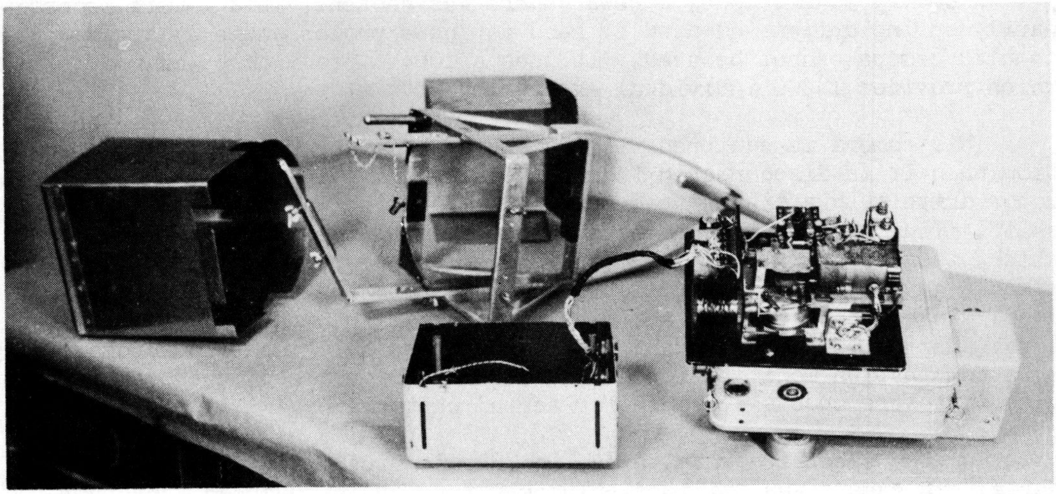

Figure 1. Camera and frame disassembled.

A 24-volt, gun sight lamp was used to provide a flash of light after the film was advanced. The lamp current was controlled by relays connected to the exposure solenoid while sufficient time for the filament to heat was furnished by the use of an eddy-current delay relay.

INTERVALOMETER

A special intervalometer was designed which would allow the use of selected time intervals such that, if desired, miles per hour could be read directly in terms of feet on the highway. This intervalometer, however, was not built (again due to lack of time and manpower).

Therefore, for try-out purposes a surplus Air Force B2 intervalometer was adapted for use with the camera.

SELF-ORIENTING MOUNT

For most effective results, it was planned to develop a self-orienting pod in which the camera would be mounted and which would keep the camera aimed vertically downward at all times. This is necessary if the plane is to make turns and follow the highway without breaking the sequence of the pictures of the highway. It was possible, of course, to fly straight and level as is done in the conventional aerial mapping procedures, but experience showed that, at the altitude required, pictures will miss the highway as it curves. Most highways will be found to include considerable

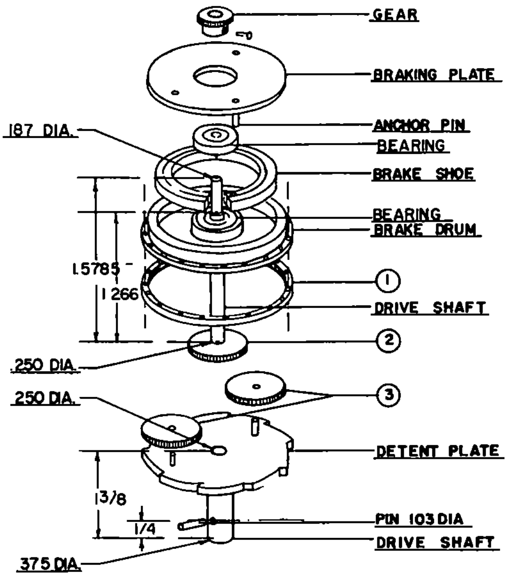

Figure 2. Exploded view of planetary drive.

curvature. Flight at higher altitude would preclude the use of a small film, such as the 35-mm. movie film used by the camera described, while the use of such film is necessary in order to keep the cost within reason.

It proved impossible to build a self-orienting pod within the limits of time and manpower available, and therefore a frame and cowling was built to allow use of the camera by hand.

FRAME AND COWLING

This frame was fashioned out of aluminum strip and designed so that the camera could be clamped into the frame by means of thumb screws and rubber pressure pads. A cowling was also built to reduce air resistance and to protect the camera from oil or other flying particles. A steel pin and rubber shock pad were provided on one corner of the assembly so that it could be set inside the steel bar which forms the cabin step on the particular airplane used.

TEST FLIGHTS

After suitable checks of the camera on the ground, a series of four flights was made to test the practicability of the camera in the air. Flights were made during the last two weeks of July 1951 and pictures were taken over Los Angeles freeways at an altitude between 3,500 and 4,000 ft. A 1947 Piper Super-Cruiser was used with the door removed.

The camera was loaded, locked into its frame and cowling and checked for operation before take-off. For safety in handling, the camera was secured to a cross member of the baggage compartment supporting frame by a doubled rope just long enough to allow the camera to be set into position and held vertically against the slip stream.

No actual sighting mechanism was used, but rough sighting was done by means of bolts on the camera frame and cowling. Since very little aiming of the camera was possible to keep it vertical in turns, the attempt was made to fly straight and level over as much of each freeway as possible, the pilot sighting by means of a spot on the landing gear below him.

RESULTS OBTAINED

After eliminating certain mechanical difficulties with the camera in the early flights, two 100-ft. rolls of satisfactory pictures were obtained. These were sharp enough to project up to a width of 4 ft. or more, without losing definition (a magnification of 40 diameters). In fact, greater magnification might be possible.

Standard 35-mm. black-and-white Eastman Plus X film was used with an Aero No. 2 Eastman filter. Figure 3 shows a sequence selected from one of these films.

Due to the curvature of the highways and the lack of an automatic orienting device for the camera, it was possible to obtain only relatively short stretches of highway continuously in the picture before a turn was necessary. It was soon demonstrated that most highways contain more curvature than one often realizes.

The experimental flights indicated that a sighting device for the camera, as well as one for the pilot, is needed. The rough sighting method used often resulted in losing part of the highway just off the edge of the picture.

The flights were made at between 3,500 and 4,000 ft. and 60 to 65 mph. indicated air speed. This altitude was convenient, due to certain defense zone flight regulations, but proved to be somewhat lower than might be desirable. Somewhat greater altitude could have been used without too great loss of detail in the picture and would have given a longer record of a given group of vehicles.

The adapted B2 intervalometer was not entirely satisfactory, since the intervals were not sufficiently accurate in the minimum time range (2 and 3 sec.) and since it usually resulted in taking more than one frame at each exposure, thus using more film than necessary.

Even so, a sufficient number of useful sequences were obtained to show that the method is practical for the study of many phases of driver-vehicle behavior. For example, in Figure 3 the entry of vehicles onto a high-speed freeway from an interchange ramp can be followed. It will be seen that the drivers of incoming vehicles were merging smoothly with the traffic stream. The pictures also show the contrast between a freeway intersection and an ordinary street intersection under heavy-traffic conditions.

Since the pictures sweep over the highway more or less with one group of vehicles but against those coming in the opposite direction, a new type of sampling of traffic is obtained. Thus, instead of the behavior at a single location only, driver activities, traffic volumes, etc., can be sampled over a number of miles of streets and highways with ease.

SUMMARY AND CONCLUSIONS

1. Aerial pictures of highway traffic were obtained using a reasonably priced, 35-mm. movie camera modified for the purpose. Satisfactory definition was obtained by projecting the pictures at a magnification up to 40 diameters with a 2- by 2-in.-film slide projector. The use of such a camera will allow 1,200 exposures without reloading, thus making it possible to follow driver behavior on the highway without excessive cost for film.

Figure 3. Sequence showing traffic at a street intersection and entering a freeway interchange.

2. Pictures taken from a light airplane at between 3,500 and 4,000 ft. obtained information on driver behavior which would be very difficult to obtain in any other way. Somewhat greater altitude could have been used to advantage. Rental of a light plane is not excessive in cost, relatively speaking.

3. Further work is necessary on aiming and sighting mechanism for obtaining vertical pictures while following the curvature of highways, and a special intervalometer must be developed for the modified camera. The ordinary aerial-mapping technique of taking photos in straight and level flight is not satisfactory for the study of driver behavior if cost is to be held within the range of most university and traffic research organizations.

4. Acquaintance with the usual problems of air photography and some practice on the part of the pilot in maintaining a track vertically over the highway will be necessary even when these devices are available. However, the method has been shown to be both feasible and promising for the study of driving behavior on the highways.

5. A new type of traffic sample is provided by the sweeping characteristic of the pictures.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge the assistance of Gordon Brittle, as air photographer while the senior author piloted the light plane, and the backing of Dean L. M. K. Boelter of the Department of Engineering, and of J. H. Mathewson, assistant director of the ITTE, which made possible the arrangements for tryout in the air.

REFERENCES

1. Johnson, A. N., Maryland aerial survey of highway traffic between Baltimore and Washington, Proc. Highway Research Board, 1928, 8, 106-115.

2. Forbes, T. W., Methods of measuring judgment and perception time in passing on the highway. Proc. Highway Research Board, 1939, 18, 100-112.

3. Greenshields, Bruce D., The potential use of aerial photographs in traffic analysis. Proc. Highway Research Board, 1947, 27, 291-297.

4. Bennett, Francis P. and Owen, Thomas. Photographic traffic surveying from a helicopter. A report of an experiment. City Plan Board, Dayton, Ohio, 1948 (mimeo.)

5. Terry, Deane S. A method of determining speed and volume by aerial photography. Unpublished Thesis. Yale Bureau of Highway Traffic, New Haven, 1949.

The Highway Research Board is organized under the auspices of the Division of Engineering and Industrial Research of the National Research Council to provide a clearinghouse for highway research activities and information. The National Research Council is the operating agency of the National Academy of Sciences, a private organization of eminent American scientists chartered in 1863 (under a special act of Congress) to "investigate, examine, experiment, and report on any subject of science or art."