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					Total number of operators
	0	1	2	3	4	
	%	%	%	%	%	%
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.

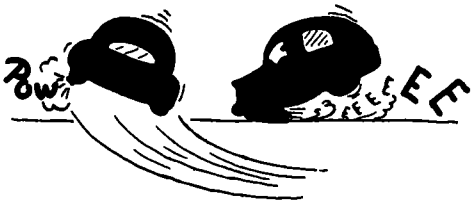
Each scale was divided into 5 sections (3 scale-points each) with a verbal definition of each section printed therein. On 10 of the 11 scales, cartoons designed to illustrate the intent of the lead question were printed at the "good" and "bad" ends of the continuum. The 11 scales were printed in booklet 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.



RATING		HOW WELL DOES HE REACT TO SUDDEN CHANGES OF TRAFFIC CONDITIONS?	RATING
!	1		A
!	2	Has no trouble at all making sudden changes.	B
!!	3		C
!!!!	4		D
!!!!!!!	5	Makes sudden changes easily.	E
!!!!!!!!!!	6		F
!!!!!!!!!!!!	7		G
!!!!!!!!!!!!!!	8	Makes sudden changes about as well as most drivers.	H
!!!!!!!!!!!!	9		I
!!!!!!!!!!	10		J
!!!!!!!	11	Has more trouble than most drivers in making sudden changes.	K
!!!!	12		L
!!	13		M
!	14	Often cannot make sudden changes.	N
!	15		O
			P
			Q
			R
			S
			T
			U
			V
			W
			X
			Y
			Z

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	h^2/	Scale	I	II	III	h^2/
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.

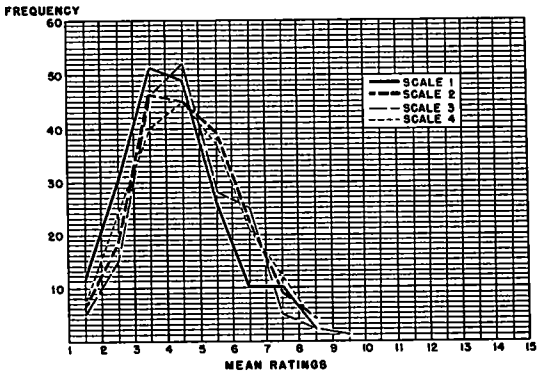


Figure 2. Distributions of mean ratings of 189 drivers rated by associates.

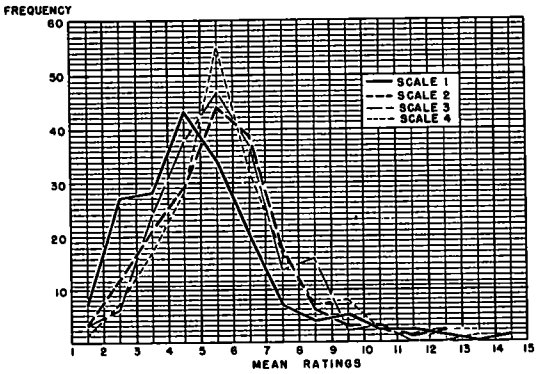


Figure 3. Distributions of mean ratings of 181 drivers rated by supervisors.

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 driv-

ing, 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"		Percent Checking "I Know I Can Rate on This"	
		A*	S*	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

	Item Number														
Scale 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.