Development of a Criterion for Driving Performance

A. R. LAUER and VIRTUS W. SUHR, Iowa State College, and
EARL ALLGAIER, American Automobile Association

●DURING the several years that driving researches have been in progress, experimenters have felt the need for a sound criterion of driving ability. Different investigators have used various types of criteria. None have been found completely satisfactory.

The ordinary road test as a criterion has been found by Lauer and others (2) to have low reliability. Even those who originally developed this test have never published anything which indicates the statistical reliability of the test. Neither has anything been done to show its validity.

The Adjutant General's Office of the Army has done a great deal to develop various types of criteria for the Armed Forces. One of the studies which has been published describes a well developed criterion of driving ability. (3) It is based on ratings by associates and superiors. While satisfactory for the purpose devised it would be rather impractical and difficult to administer under civilian conditions due to the lack of a sufficient number of persons familiar with the driver's performance at the wheel who could rate the driver.

The criterion of reported accidents has been found quite unreliable. Even over successive periods Johnson (1) found correlations as low as .30 for the same drivers in two successive epochs. Various explanations may be given. First, there is very little chance of being caught in a violation or accident, and second, there seems to be a psychological effect which follows being involved in an accident. The precise effect seems to be differential. For most drivers it has a deterrent effect. For a few persons it is indicative of future behavior with similar results. Intercorrelations of rating scales, as well as laboratory devices designed for measuring certain aspects of driver performance, were made. The hypothesis set up for testing may be stated as follows: patterns of behavior relating to good driving performance are measurable.

METHOD AND PROCEDURE

This is a four-phase study of driver performance designed to establish a basis or criterion of driving ability. Each subject was given a simulated driving test in the laboratory under controlled conditions. The second test was that of driving an instrumented car over an 8-mile standard route. A tachograph record was obtained for each subject while the trip was being made. The road driving performance was rated by means of the Roger-Lauer Scale. Three hundred forty-nine subjects were used. They included commercial drivers, lay drivers, and students just completing a driver education course.

The Roger-Lauer Scale

This is a paper and pencil rating scale first developed in 1938. It is divided into two parts. Each part is scored separately and the two combined for a total score.

Part one covers twelve behavior categories assumed to be largely inherent. There are five descriptive phrases in each category. The rater checks the phrase which he thinks is most characteristic of the rates. Each phrase has a numerical weighting. The sum of the numerical weightings constitute the score.

Part two is concerned with the degree of skill shown in performance of certain functions deemed basic to efficient automobile driving. It consists of 15 items which are rated on a seven-point scale. The sum of the scale point values constitutes the score.

This scale was originally reported to have a reliability of the order of .90. Split-

TABLE 1

RELIABILITY OF RATINGS ON ROGER-LAUER SCALE

	Students[a]		Experienced Drivers[b]		Total Group[c]	
	r_{12}	r^d	r_{12}	r^d	r_{12}	r^d
Part 1	.76	.86	.57	.73	—	—
Part 2	.76	.86	.72	.84	—	—
Total score	—	—	—	—	.85	.92

[a] Based on 231 cases
[b] Based on 118 cases
[c] Based on 349 cases
[d] Estimated full-length reliability using the Spearman-Brown formula

TABLE 2

INTERCORRELATIONS OF TACHOGRAPH AND ROGER-LAUER SCALE DATA

	1	2	3	4	5
1. Trip time	—	-.6920	-.6590	.5640	-.2690
2. Modal speed		—	.8290	-.4970	.2000
3. Maximum speed			—	-.5050	.1350
4. Number of fluctuations				—	-.0003
5. Roger-Lauer Scale					—

A correlation of the Tachograph measurements with the Roger-Lauer Scale rating yielded a multiple R of .3375 which shrank to .3178 when corrected for the number of cases and number of variables.

On a rational basis from the magnitude of the betas it would seem that (1), (2), and (4) of this matrix should be considered in the final evaluation phase of this study.

TABLE 3

RELIABILITY OF DIFFERENT SCORES ON THE AUTO TRAINER[a]

	r_{12}	r^b on sum of test-retest scores
Steering (contacts)	.52	.68
Total time (minutes)	.75	.85
Response time (milliseconds)	.48	.66
Movements (total recorded)	.48	.66
Errors (practices violated and mistakes made)	.73	.84

[a] Based on 150 cases at Iowa State College
[b] Estimated reliability of a test-retest score using the Spearman-Brown Formula—scores added together.

half reliabilities were computed from the ratings made in this study. The resulting coefficients for part one, part two, and total score are shown in Table 1.

Tachograph

The Sangamo Model AA 12-hour Tachograph was used. In addition to containing a speedometer, odometer, and clock, this instrument makes a graphic recording of fluctuations in speed, total trip time, and miles traveled. A sample tachograph chart is shown in Figure 1.

An analysis of each trip was made by means of the tachograph chart analyzer shown in Figure 2. Trip time, modal speed, maximum speed and number of fluctuations were determined.

The intercorrelations of the various factors measured by the tachograph together with their correlations with the Roger-Lauer Scale ratings are shown in Table 2.

A multiple correlation of .3375 was obtained between the four tachograph variables

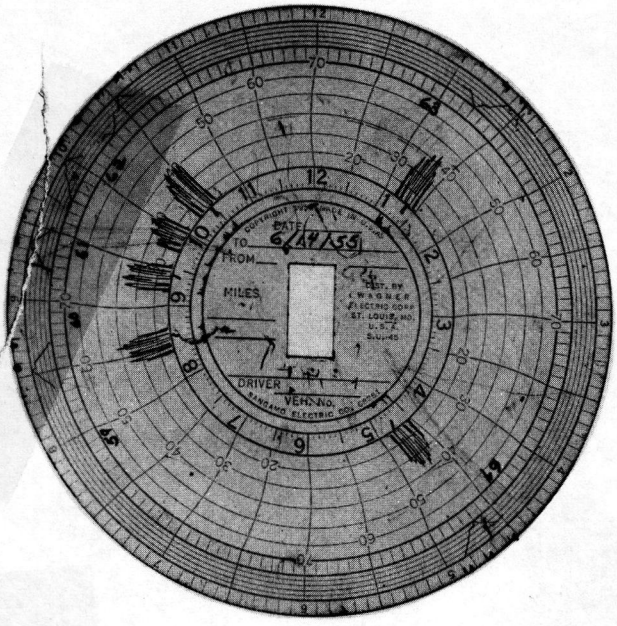


Figure 1. Tachograph Chart.

Record made on the Sangamo 12-hour Tachograph. Six short drives are shown. The following factors are measured by the Tachograph Chart Analyzer shown below: maximum speed, modal speed, fluctuations in speed, total trip time, and miles traveled.

A multiple R of .33 was obtained with the Roger-Lauer Rating Scale as the criterion.

Other factors could be measured as thought advisable. It would appear that a great deal of valuable information about drivers could be obtained from a scientific analysis of their records.

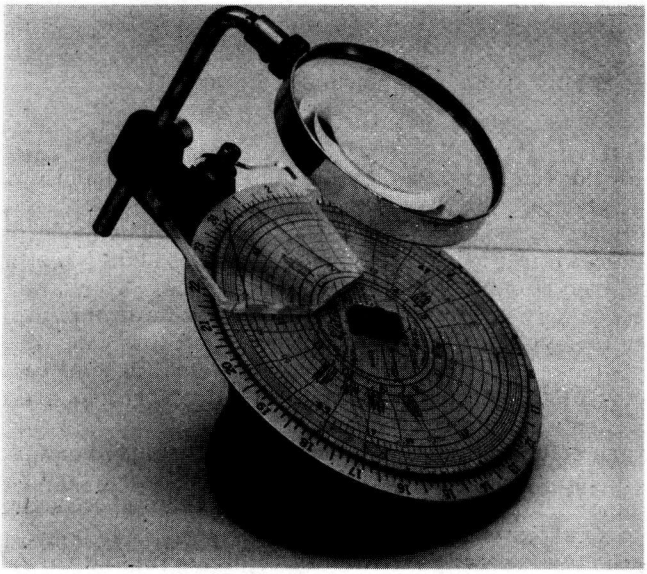


Figure 2. Tachograph Chart Analyzer.

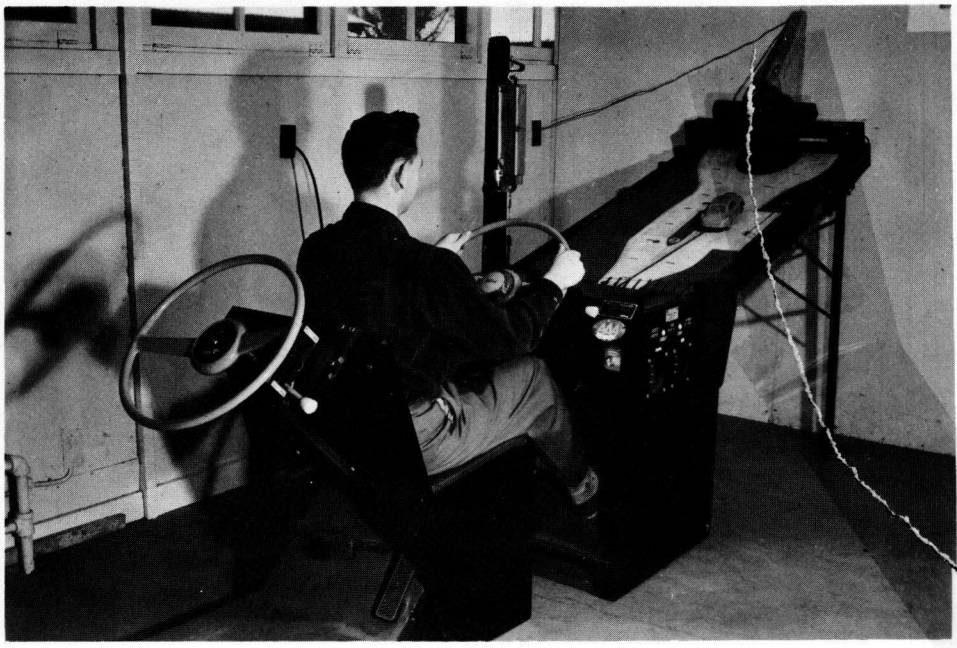


Figure 3. Model B Auto Trainer.

AAA Model B Auto Trainer is used as a laboratory test and training device. The various factors measured are (1) steering efficiency, (2) errors made in manipulation, (3) reaction or response time to a red signal light, (4) extent of movement made and (5) time for the trip.

As a testing device it correlates .45 with road driving as rated by the Roger-Lauer Scale. This scale has a reliability of .92.

As a training device it can be used singly or in any reasonable number of units. For small schools, multiples of four units are recommended. The first four lessons in fundamentals of driving can be taught much more economically than in a car. Extra practice can be given without risk at the student's convenience.

and the Roger-Lauer Scale rating. The shrunken R was .3178.

Auto Trainer

This is a laboratory device developed by the American Automobile Association for purposes of driver instruction. Full-size automobile controls are used to guide a miniature car around a traveling roadway simulating driving. The Model B Auto Trainer used in this study is shown in Figure 3. It may also be used as a testing device.

The device is designed so as to yield several subscores. They are (1) steering efficiency or the ability to stay on the road, (2) response time to traffic lights as presented, (3) errors in manipulation or such failures as not following directions and road signs, (4) movements made in braking, shifting, etc., and (5) total time for the trip of a given number of revolutions of the roadway belt. All recordings are made automatically by a battery of electric counters.

The conditions were standard for all subjects. They included use of clutch, brake and accelerator, steering wheel, driving forward and backward, and parking. Observation of signs and other features of the device simulate actual road-driving conditions and are controlled to some extent.

Table 3 contains the test-retest reliability coefficients for the various scores on the Auto Trainer.

The intercorrelations of the Auto Trainer scores are listed in Table 4. Correlations with the Roger-Lauer Scale ratings are also included in this table. The six Auto Trainer scores yielded a multiple correlation of .4503 with the Roger-Lauer Scale rating. The corrected R shrunk to .4289.

TABLE 4

INTERCORRELATIONS OF AUTO TRAINER SCORES AND ROGER-LAUER SCALE RATING

	1	2	3	4	5	6	7
1. Steering	—	-.3120	.1050	-.2960	.1710	.2080	.1480
2. Errors		—	.2150	.1840	.1780	-.2270	-.2170
3. Movements			—	-.0090	.4350	-.1600	-.1360
4. Observation time				—	.0520	-.1140	-.0400
5. Total trip time					—	-.2540	-.3610
6. Hand brake pressure						—	.2660
7. Roger-Lauer Scale							—

A correlation of the Auto Trainer scores with the Roger-Lauer Scale rating yielded a multiple R of .4503 which shrank to .4289 when corrected for the number of cases and number of variables.

On a rational basis from the magnitude of the betas it would seem that (1), (2), (5), and (6) of this matrix should be considered in the final evaluation phase of this study.

Car Instrumentation

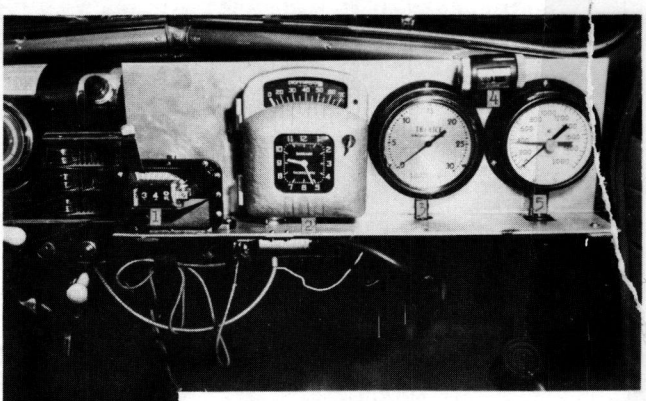
An Oldsmobile hydramatic drive, four-door sedan equipped with instrumentation designed to objectively measure driver performance with respect to certain factors deemed relevant to efficient operation of a motor vehicle was used for the road tests.

On a panel just back of the dashboard at the right side the following instruments are mounted from left to right as shown in Figure 4.

1. Revco reduction gear and counter. It is used to integrate the total amount of steering-wheel movement made by the driver. The counter is attached to the reduction gear so as to make a continuous recording of the steering wheel movements in both directions. The reset counter is set at zero at the beginning of each trip. At the end of the trip the numerical reading is recorded as steering-movement score.
2. Sangamo Model AA Tachograph previously described.
3. Trerice vacuum gauge. The gauge is attached to a pressure chamber which is set to activate a counter whenever the carburetor vacuum pressure is equal to the setting. In this way a recording of accelerator movements is obtained.
4. Gasoline meter. A McCulloch gasoline meter was used to measure gasoline consumption in $\frac{1}{100}$ths of a gallon. The meter can be reset to begin accumulating from zero at the beginning of each trip.
5. Trerice hydraulic pressure gauge. Maximum pressure in pounds made on the brake pedal is measured and retained on this instrument by a special hand.

Other instruments are located in the rear seat area of the experimental car. They are shown in Figures 5 and 6.

6. Accelerator movement counter. The counter is mounted on a panel built infront of the rear seat rest. It is connected to the vacuum gauge in such a way as to be activated every time the indicator hand of the gauge fluctuates above or below a preset reading.
7. Brake movement counter. This recording device is mounted just below the accelerator movement counter. It is connected to the stop light circuit so that it is activated every time the brake pedal is depressed thus recording brake applications.
8. Jerk recorder. The jerk recorder is housed in a steel cabinet measuring $4\frac{1}{2}$ in. x $7\frac{1}{2}$ in. x 5 in. high. A $\frac{1}{2}$ in. steel shaft runs through the center of the cabinet near the top. Two pendulums each $4\frac{1}{2}$ in. long are mounted to swing freely on this shaft.



INSTRUMENT PANEL. 1. Steering-wheel movement 2. Tachograph 3. Vacuum gauge 4. Gasoline meter 5. Hydraulic brake pressure gauge

Figure 4.

TOP. Instrumentation on dash of the experimental test car for drivers. (1) Revco reduction gear with attached counter measures the amount of steering done over a given route. (2) The Sangamo Tachograph yields a number of measures including fluctuations in speed, total trip time, and miles traveled. (3) Accelerator movements are measured by a Trerice vacuum gauge. (4) The McCulloch gasoline meter gives the gasoline consumption in hundredths of a gallon. (5) Braking is recorded by the Trerice pressure gauge. The maximum indicator hand reading is recorded. Electric counters are also attached to (3) and (5) to give accumulative records.

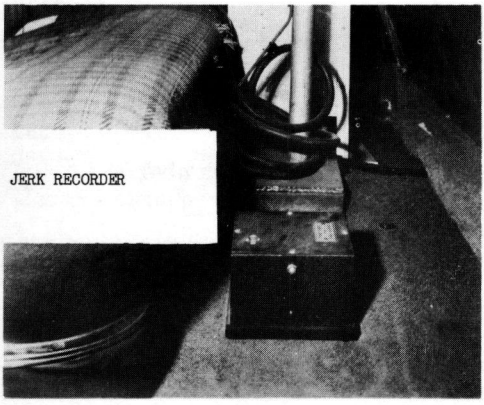


Figure 5.

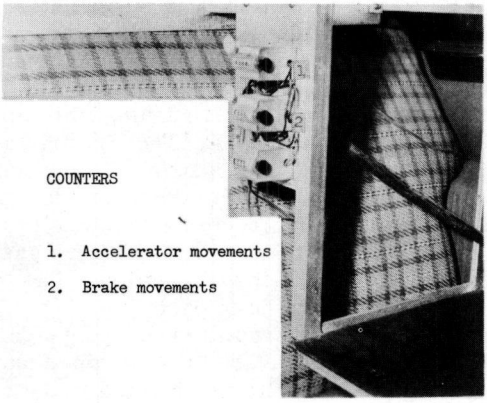


Figure 6.

On the center of the shaft between the two pendulums is an assembly consisting of two brass ratchets each with 60 teeth and a sprocket wheel with 20 teeth.

When a stop is made one pendulum swings forward turning one of the ratchets by means of a small pawl. The sprocket wheel turns with the ratchet and operates a mechanical counter. Each 18 degrees of rotation of the pendulum counts one unit on the counter. A movement of six deg. causes the counter to record ⅓ unit. The other pendulum is arranged so that it moves when a sudden start is made and operates in a similar manner but with different settings for sensitivity.

Two of these instruments are used. One is placed lengthwise with the car so as to record sudden stops or quick starts. The other is placed crosswise with the car so as to measure sway from side to side. It is set to register slighter movements, that is, made more sensitive. Both instruments are placed on the floorboard in front of the rear seat.

Split-half reliability of the data gathered by the various instruments in the experimental car is shown in Table 5.

TABLE 5

RELIABILITY OF EXPERIMENTAL CAR INSTRUMENT SCORES[a]

	r_{12}	r^b
Steering movements	.54	.70
Gasoline consumption	.62	.77
Accelerator movements	.70	.83
Brake movements	.42	.59
Brake pressure	.48	.65
Jerk recorder (lengthwise)	.70	.83
Jerk recorder (crosswise)	.56	.72

[a] Based on 349 cases
[b] Estimated full-length reliability using the Spearman-Brown formula

TABLE 6

INTERCORRELATIONS OF EXPERIMENTAL CAR, INSTRUMENT SCORES AND ROGER-LAUER SCALE RATING

	1	2	3	4	5	6	7	8
1. Gasoline consumption	——	.3220	.1970	.0780	.1600	.0080	.0090	-.1030
2. Steering movements		——	.2630	.0710	.3870	.1050	.1730	-.5210
3. Brake pressure			——	.1490	.0780	.1220	.1620	-.1940
4. Brake movements				——	.2410	.1720	-.1220	.0510
5. Accelerator movements					——	.2240	.0690	-.2520
6. Jerk recorder (lengthwise)						——	.3760	-.2380
7. Jerk recorder (crosswise)							——	-.1760
8. Roger-Lauer Scale								——

A correlation of the Experimental Car instrument scores with the Roger-Lauer Scale rating yielded a multiple R of .5741 which shrank to .5603 when corrected for the number of cases and number of variables.

On a rational basis from the magnitude of the betas it would seem that (2), (4), and (6) of this matrix should be considered in the final examination phase of this study.

Intercorrelation of the scores obtained from the experimental car instruments together with their correlations with the Roger-Lauer Scale rating are presented in Table 6.

A multiple correlation of the instrument scores with the Roger-Lauer Scale rating was computed. R = .5741. This shrinks to .5603.

DISCUSSION

This is the first part of an extended study which is being made in an effort to develop an objective criterion of driving performance. The purpose of this portion of the study was to determine the reliability and validity of the various objective separate measures and to select the ones which seem to be most worthy to be included in the second part of the study which is to be reported later.

The Roger-Lauer Scale ratings were used as the primary criterion against which to evaluate the potential predictive value of the several objective scores. This scale was selected because it has sufficient reliability for individual use and provides an immediate criterion. It also samples behavior patterns as well as developed skills.

Two of the Auto Trainer subscores, namely total time and errors, seemed to possess sufficient reliability and validity to be retained for further study. These two measures with hand brake pressure correlated the highest with the Roger-Lauer Scale rating. Reliability of hand brake pressure was not computed. It was used more as an auxiliary measure in this study. From previous studies it would appear that steering movements should be retained for further evaluation.

Three of the tachograph variables seemed promising. They are trip time, modal speed, and fluctuations in speed. Modal speed correlated considerably higher with the Roger-Lauer Scale rating than did maximum speed.

Most of the instruments in the experimental car seemed to measure with sufficient consistency to merit their retention in the second part of the study. Accelerator movements and smoothness as measured by the jerk recorded as measuring lengthwise thrust of the car showed highest reliability.

The measurements from the four phases of the study showing satisfactory performance are to be combined into a multiple correlation with road driving performance rating in the second part of this study. The factors making the most significant contribution to the multiple R are to be used in development of an objective criterion of driving performance ability. From the data available it would appear that with these techniques combined it may be possible to measure from 70-90 percent of the variance in driving ability.

CONCLUSIONS

Considering the characteristics of the sample employed, number of subjects, and the nature of the results obtained, the following conclusions may be tentatively drawn from this study:

1. The Auto Trainer yields three subscores, namely, total time, errors, and steering which are sufficiently substantial to justify their use in further research in driver evaluation studies.
2. Analysis of tachograph records will give valuable information on driving ability. Further study needs be made of possible derived scores.
3. The AAA Jerk Recorder will reliably measure smoothness of movement when placed lengthwise with the car.
4. The better performing driver holds the wheel steadier, i.e., turns the wheel less, uses less gasoline, works the accelerator less and is less severe on the brake than the poorer performing driver.
5. Certain factors related to road driving performance can be substantially measured by means of scores objectively obtained while the driver is performing the task. Supplementary measures made by simulated driving devices will increase the predictive value of a battery designed for this purpose.
6. The hypothesis set up for testing—that driving performance can be objectively measured—is affirmed within reasonable limits of error.

ACKNOWLEDGMENT

This study was made possible through a grant for driving research from the American Automobile Association.

REFERENCES

1. Johnson, H. M. The detection and treatment of accident-prone drivers. Psychological Bulletin, 1946, 43 (no. 6), pp. 489-532.
2. Lauer, A. R. and others. Selection tests for army operators of light and heavy motor vehicles. U. S. Army, PRB Technical Research Report 1081. 1953, p. 22.
3. Uhlaner, J. E., Goldstein, Leon G., Van Steenberg, N. J. Development of criteria of safe motor-vehicle operation. Highway Research Board Bulletin 60. 1952, p. 16.