Final Report of Study on the Effect of A Periodic Refreshment Pause on Simulated-Automobile Driving Performance Efficiency

VIRTUS W. SUHR, Iowa State College

During a 3-year period, 56 drivers were individually taken through an experimental 15-hr cycle. The first phase was to familiarize the subjects with the equipment and to minimize any practice effect that might be reflected. Because a pilot study showed that for variables of this type maximum performance was reached within the first 3 hr of driving, the practice period was limited to 3 hr behind the wheel.

The second phase studied driving efficiency over a 6-hr period of simulated automobile driving. An attempt was made to determine the performance curve for variables amenable to continuous measurement. A series of psychological, physiological, psychometric, and psychophysical tests was administered both before and after the driving period to detect any changes that occurred in performance on certain factors relevant to safe driving. The effect of periodic tea pauses on the onset and extent of work decrement was noted. A personal evaluation from each participant was compared with driving performance and test results.

The principal findings are as follows: (a) work decrement begins within the first 2 hr of simulated automobile driving; (b) tea pause prolongs the onset and reduces the work decrement resulting from a prolonged period of simulated automobile driving; (c) the effect of a tea pause can be detected most readily by measurements made while the subject is actually performing; and (d) drivers either cannot or do not accurately evaluate their own level of driving efficiency.

THE PROBLEM of exhaustion and drowsiness among drivers has been of concern for sometime to those working in the field of traffic safety. As early as 1935 the National Safety Council (1) reported urgent need for scientific research to determine what effect continuous driving, long hours without sleep, and various mechanical aspects of vehicles, have on fatigue and to ascertain the recuperative value of various rest periods. Such research should also bring out, if possible, what degree of fatigue is dangerous and how different individuals vary in susceptibility.

In 1936 Ryan and Warner (2) completed a study on the effect of automobile driving on the reactions of the driver. They concluded that the tendency of long automobile drives is to produce a loss of effectiveness of certain sensory discriminations, association processes and motor reactions similar to those required in driving. These observations suggest that the effect of a long automobile drive may render a driver temporarily prone to accidents.

The present investigation is a study of driving efficiency. Experimentation covered a period of approximately 15 hr for each subject. The primary objective was to determine the effect of a periodic refreshment pause on performance efficiency with respect to a laboratory-driving task. The nature and extent of work decrement for driving performance were also studied.

The experiment was divided into two phases. The first or practice phase was introduced to familiarize the subjects with the equipment and to minimize any practice effect that might be reflected in performance. Because it was shown in a pilot study that for variables of this type maximum performance was reached within the first 3 hr of driving, the practice phase was limited to 3 hr behind the wheel.

In the second phase the driving time was lengthened to a 6-hr period of performance on the simulated-automobile driving task. An attempt was made to determine the nature of the performance curve for certain variables amenable to continuous measurement.

A series of psychological, physiological, psychometric and psychophysical tests was administered both before and after the driving period in an effort to detect any changes that may have occurred in performance on certain factors deemed relevant to safe driving. Periodic refreshment pauses were introduced and the effect on the onset and extent of work decrement was noted. A personal evaluation was obtained from each participant about subjective feelings which could be compared with driving performance and test results.

METHOD AND PROCEDURE

The method was that of an experimentally controlled study using comparison groups matched as nearly as possible with respect to sex, age, and driving experience. The groups were balanced approximately in the ratio of seven men to three women, which is commensurate with that found in the driving population of the United States. The two groups were used in a simulated-automobile driving performance task.

During the first phase of the study the control group, henceforth designated as the no-pause group, drove for three consecutive hours without rest. The experimental group, hereafter referred to as the refreshment-pause group, spent the same amount of time behind the wheel but were given a 15-min rest pause with tea and additives as desired every 1½ hr during the experimental period.

In the second phase of the study the no-pause group drove for six consecutive hours. The refreshment-pause group also drove for 6 hr, but was given a periodic rest pause with tea as described every 1½ hr.

The first phase of this investigation was designed to test the hypothesis that performance of a task of the type studied is subject to improvement with practice which may be affected by refreshment pauses and basic efficiency changes in the organism.

The primary hypothesis set up for testing in the second phase of the study may be stated in the null form as follows: A periodic refreshment pause (namely, a 15-min rest period every 1½ hr during which time tea as described is served) has no effect on the simulated-automobile driving task performance of automobile operators.

A corollary hypothesis, that work decrement as a result of an extended period of simulated-automobile driving will not be reflected in performance on laboratory tests under the conditions stated, was also tested.

Fifty-six drivers, including 38 men and 18 women, served as subjects. They were recruited from among lay drivers in the vicinity of Iowa State College and other Central Iowa communities. The selection was restricted to drivers having had at least 3 yr or 10,000 mi of driving experience. A second restriction was imposed by the fact that the drivers be willing and able to devote a total of approximately 15 hr to participation in the study. Compensation was mostly at their regular hourly rate of pay. The selection was in order of availability of subjects as arrangements could be made.

The drivers ranged in age from 18 to 66 years. The median age was 24 years for the men and 25 years for the women, the mean being 28.75 years and 26.40 years for men and women, respectively.

Driving experience in terms of years driven ranged from 2 to 38. The median was 8 years.

Miles driven, as estimated by the subjects, ranged from 2,000 to 564,000, with medians of 70,000 and 16,000 for the men and women, respectively.

The experimental procedure began by having each subject fill out a preliminary information blank containing personal data and self-evaluation items. As soon as the preliminary information blank was completed a series of evaluation tests was administered in the following order:

1. Blood pressure. The Tycos self-recording sphygmomanometer was used for measuring blood pressure. This instrument is particularly adapted for use in this type

of study because it makes a graphic recording which increases the objectivity of measurements obtained.

2. Steadiness test. A stylus 1/8-in. in diameter is moved down between two brass strips which are 3/8-in. apart at the top and 1/8-in. apart at the bottom. When either plate is touched a light flashes and the trial ends. The score is read from a calibrated scale on one of the plates. A series of ten trials, alternating hands each trial, constitutes the test.

3. Serial choice reaction time. The subject is seated with the right foot placed on a break-type switch adjacent to a simulated brake pedal and is instructed to hold the right foot on the switch just as though pressing the accelerator of an automobile.

Green, amber, and red stimulus lights are presented in random order. The subject is instructed to respond only to the red light; that is, as soon as the red light appears he is to move the right foot from the switch and place it on the brake pedal as quickly as possible.

The apparatus records only the reaction time to the red light. False or wrong reactions, such as responding to a green light, are merely counted. The test continues until the red light is presented 25 times. Several amber and green lights are interspersed as distraction stimuli in the series of 50 presentations. The number was constant for each subject.

4. Gross coordination. This is measured by a device developed at the Driving Research Laboratory, Iowa State College, for use with Army drivers (3). A tilting-table maze is controlled by means of two levers approximately 43 1/2 in. long. One lever tilts the table top upward or downward from front to back, the other tilts it in a similar manner from side to side. A steel ball can be guided around the lanes of the maze by manipulating the levers. At various places along the course are located 1-in. holes, through which the ball will drop if the levers are not manipulated properly to maneuver it around them. The object is to guide the ball through the maze without its falling into any of the holes. When the ball falls through, a trial is completed.

The holes are numbered progressively, so that the farther the ball advances around the maze before it falls through a hole, the higher the score, as the number of the hole determines the score value for the trial. In this experiment each subject was given five trials and the mean of the scores was recorded.

5. Grip endurance. (Not used in Phase I). The apparatus used was a Smedley hand dynamometer with pneumatic plunger attached to a tambour mounted on a Weiss-Renshaw polygraph, which in turn made an inked recording on ruled paper. The subject was asked to take the dynamometer in the preferred hand and grip it as hard as possible for 1 min. The percentage of loss from the original level attained was taken as the score.

6. Card sorting. (Not used in Phase I). The equipment consisted of a deck of Rook cards and four small boxes with one of the following colors printed on the front of each: yellow, red, green, or black. The cards were shuffled and handed, face down, to the subject with the instructions to turn the deck over, look at the top card, state aloud the color, then place it in the proper box. The subject was given only one chance for each card; thus, if a card were placed in the wrong box, it remained there. The object was to see how rapidly the cards could be sorted. The number of errors was also recorded.

7. Mental addition. (Not used in Phase I). Twenty addition problems, each composed of five two-digit numbers, constituted the test. One minute was allowed for computation. The number of problems attempted and number of errors were recorded. Alternate forms were used for the pre-testing and post-testing sessions.

8. Efficiency in observing, or attention to detail. This is a paper-and-pencil test consisting of several rows of the same letter. From one to four other letters of near-identical design were inserted in some of the rows. Two minutes were allowed for counting the number of odd letters in the several rows. Rows attempted and errors were recorded.

9. Galvanic skin response, pulse, and respiration. A Stoelting No. 22496 deceptograph was used for obtaining these measurements. The subject is seated comfortably in a lounging chair and told to relax as much as possible. A pneumatic cuff is placed around one wrist and inflated sufficiently to bring out the pulse beat.

A pneumograph is fastened around the chest tight enough to stretch and contract as the subject breathes. A finger electrode is attached with electrocardiograph jelly to the middle finger of each hand in order to obtain a measurement of skin resistance. A graphic record of pulse, respiration, and galvanic skin responses was obtained for a period of 1 min for each measurement before and after driving.

After the preliminary testing each subject in the refreshment-pause group was served tea. Subjects in the no-pause group went directly into the simulated-automobile driving task phase of the study. Upon entering the booth the subject was seated in the Drivometer, a device consisting of a mock-up landscape with a model car controlled from a control seat exactly as found in a regular automobile, and given instructions with respect to the driving task by means of a tape recorder to keep directions constant.

At the beginning of the driving period, a control test run covering a cycle of seven instructions was made on the Drivometer. During this time a red stop light was presented on five different occasions. Likewise, a train was made to emerge from a tunnel and pass in front of the driver on five other occasions at irregular intervals. The time required for the driver to do the driving task was measured by an electric time clock and recorded as total trip time. The steering score, stop light response time, train reaction time, and error time were measured and recorded each half hour.

The subjects were told to drive just as though they were on the open road as soon as the instructions ceased to appear in the aperture of the Drivometer panel. The red stop light was consistently presented five times each half hour. The train was made to emerge from the tunnel into the view of the driver on five different occasions each half hour. The steering score and number of belt revolutions were recorded for every half hour of driving. The several hours of simulated driving task has been called the intransit period.

After 1½ hr of continuous driving, the subjects in the refreshment-pause group were given a 15-min rest period and again served tea as previously described. The no-pause group drove for three and six consecutive hours in Phases I and II, respectively.

Ten minutes prior to the end of the last half hour of the intransit period, a second control test run was given. As soon as the simulated-automobile driving task was finished, the evaluation tests were administered again in the same order as previously described. The experimental cycle was completed by obtaining the driver's subjective evaluation of his level of efficiency by means of a specially designed form.

The results of Phase I were reported previously (3). Therefore, the remainder of this paper covers only Phase II of the study.

RESULTS

An analysis of covariance was made from the scores on the evaluation tests that were administered before and repeated again after the intransit period. The scores on the tests administered before the driving period were taken as the covariates. The results are presented in Table 1.

The drop in diastolic blood pressure by the no-pause groups as compared to practically equivalent measurements for the refreshment-pause group was statistically significant at the 5 percent level of confidence. The refreshment-pause group attempted more items in the test of efficiency in observing and got more items right than the no-pause group. However, they also made more errors. The difference was significant at the 1 percent level of confidence. To be sure that the difference was not spuriously inflated due to the fact that the number right was not controlled in the first analysis, a second analysis was made with the number right after the driving period taken as the covariate. The difference was still significant at the 1 percent level of confidence. A third analysis was made with the last grade of school completed taken as the covariate. The difference remained significant at the 1 percent level of confidence.

To compare the efficiency of the two treatment groups after the intransit period with respect to the variables measured during the control test run, an analysis of covariance was made with the scores on the initial performance taken as the covariates.

Regarding two of the factors, stop light response time and train reaction time, an inspection of the data revealed that some individuals took an extremely long time to respond on a few occasions. The possibility that a few extremely long intervals might spuriously inflate the group means and thereby influence the results of the analysis was considered. It was decided to adjust the data for extreme values. To do this objectively, the method devised by Grubbs (5) was followed. Both the original and adjusted data were analyzed. The results for all variables considered in the control test runs are given in Table 2. No differences of sufficient magnitude to be statistically significant were revealed.

In the intransit period the instruments used for collecting data while the subject was behind the wheel of the Drivometer were read every half hour. There were twelve readings for each of the variables considered. Group mean scores were computed for each half-hour period. The results are shown in Figures 1 through 4.

TABLE 1

MEAN RESULTS OF EVALUATION TESTS

Variable	Refreshment-Pause		No-Pause		F
	Before	After	Before	After	
Choice reaction time:					
Mean	37.300	39.429	36.536	40.321	1.374
Average variability	6.768	6.757	6.811	7.875	0.405
False attempts	0.714	0.893	1.607	1.036	0.468
Coordination	59.286	55.679	58.214	55.536	0.046
Grip endurance	35.336	35.339	34.479	34.529	0.016
Steadiness	7.871	7.536	9.093	8.686	0.319
Strength of grip	40.354	42.136	38.568	40.004	0.253
Blood pressure:					
Diastolic	67.321	67.535	64.071	62.000	4.061[1]
Systolic	120.250	123.286	122.750	121.036	0.212
Galvanic skin response	127.262	111.900	98.771	76.737	0.266
Pulse:					
Rate	77.429	72.036	76.786	70.179	0.282
Regularity	1.003	1.001	1.008	1.014	0.526
Level of oscillation	1.071	1.107	1.000	0.964	0.400
Respiration:					
Frequency	16.107	16.643	15.250	15.821	0.138
Mean I/E ratio	0.771	0.721	0.837	0.781	0.286
Mean I/E variability	0.210	0.194	0.224	0.203	0.010
Efficiency in observing:					
Number attempted	25.607	34.393	25.679	31.071	3.465
Number right	22.071	29.286	22.464	28.179	0.843
Errors	3.536	5.107	3.214	2.893	13.068[2]
Card sorting:					
Time	44.500	42.857	44.536	40.500	1.747
Errors	0.357	0.286	0.321	0.429	0.244
Mental addition:					
Number attempted	16.429	16.036	17.643	16.964	0.018
Number right	15.500	15.179	17.214	16.357	0.232
Errors	0.929	0.857	0.429	0.607	0.168

[1] Significant at the 5 percent level.
[2] Significant at the 1 percent level.

The degree to which the car was kept in the proper lane on the road was objectively recorded by means of a counter that was activated every time the miniature car passed over one of a series of copper bars in the center of the right lane of the traveling roadway. The cumulative count constituted the steering score.

Figure 1 reveals that at the end of the first half hour of driving, the scores for the two treatment groups were practically equivalent. The refreshment-pause group consistently made the higher steering scores for the remainder of the driving period.

To obtain a measure of speed, the average number of Drivometer belt revolutions per minute was recorded for each half-hour period. The performane curves of the two treatment groups were compared graphically (Fig. 2).

TABLE 2

MEAN RESULTS OF CONTROL TEST RUNS

Variable	Refreshment-Pause		No-Pause		F
	Before	After	Before	After	
Error time	0.832	0.777	1.011	0.721	1.272
Steering	106.893	103.321	129.893	98.821	2.445
Stop light response time:					
Mean:					
Original data	1.021	1.282	1.071	1.185	0.298
Adjusted for extreme scores	0.835	0.863	0.872	1.063	2.879
Average variability:					
Original data	0.500	0.680	0.580	0.454	0.882
Adjusted for extreme scores	0.157	0.163	0.148	0.279	2.539
Total trip time	4.012	3.645	4.019	3.611	0.031
Train reaction time:					
Mean:					
Original data	0.780	0.774	0.898	0.742	0.795
Adjusted for extreme scores	0.735	0.720	0.744	0.701	0.143
Average variability:					
Original data	0.475	0.567	0.649	0.444	1.249
Adjusted for extreme scores	0.359	0.360	0.399	0.367	0.037

The refreshment-pause group consistently had the greater average number of belt revolutions per minute, which indicated that they drove somewhat faster than the no-pause group. The average number of belt revolutions per minute increased during the half-hour period immediately following each refreshment pause.

The groups also were compared with respect to steering score per simulated distance traveled. As a basis for the comparison, the steering score for each half-hour period was divided by the number of belt revolutions during that period. The result, called steering efficiency (Fig. 3) reveals how well the car was kept in the proper position on the roadway.

The red stimulus light was presented on five different occasions during each half hour of the driving period. The time required for the driver to notice the red light and respond by depressing the brake pedal was recorded as stop light response time. Both the mean and average variability were computed. The results (Fig. 4) show that the no-pause group consistently took a longer time to respond to the red light than the refreshment-pause group. They also showed more variability in stop light response time than the refreshment-pause group.

A few drivers occasionally took an extremely long time to respond. Adjustment for extreme outliers was made according to the method devised by Grubbs (5). After the adjustment had been made for extreme scores the stop light response time for the two groups was practically equivalent through the third half hour. Beyond the third half hour of the driving period the no-pause group consistently took longer to respond than the refreshment-pause group.

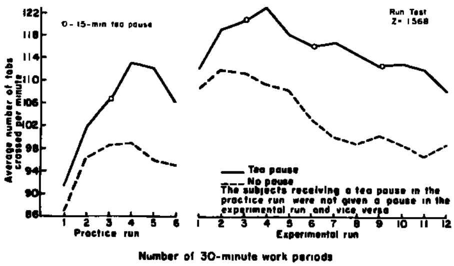


Figure 1. Results of steering tests.

The average variability for the two groups after adjustment for extreme outliers was nearly equivalent through the eighth half hour, after which the no-pause group became increasingly more variable with respect to stop light response time than the refreshment-pause group.

An electric train was caused to emerge from a tunnel into the view of the driver at five different times during each half hour of driving. The time elapsing between the emergence of the train from the tunnel and the driver's response by depressing the brake pedal was recorded as train reaction time. The mean and average variability of the two treatment groups were computed for each half hour of the intransit period. Similar computations were made after the data had been adjusted for extreme scores. The performance of the two treatment groups was practically equivalent. No substantial differences were revealed (Fig. 5).

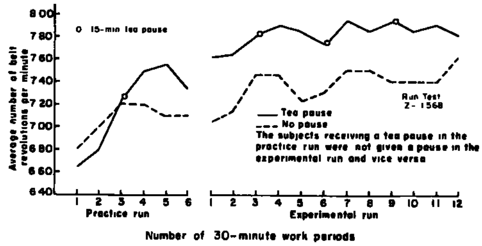


Figure 2. Speed vs driving time.

To test the significance of the differences between the treatment groups with respect to the factors studied during the driving or intransit period, the half-hour intervals were considered as split-plot trials and an analysis was made.

The resulting F-values for treatment groups, 30-min intervals, and interaction as obtained for each variable, are presented in Table 3.

The analysis revealed a highly significant F-value with respect to the number of belt revolutions (that is, simulated distance traveled) in each half hour of the driving period. This indicates a statistically significant difference in the number of belt revolutions among the 12 half-hour intervals.

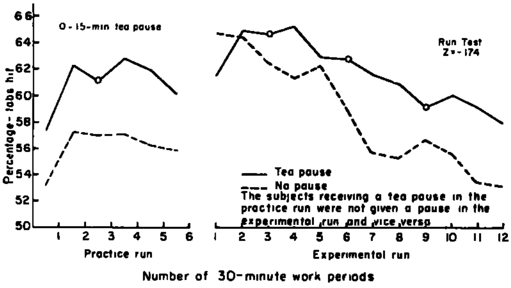
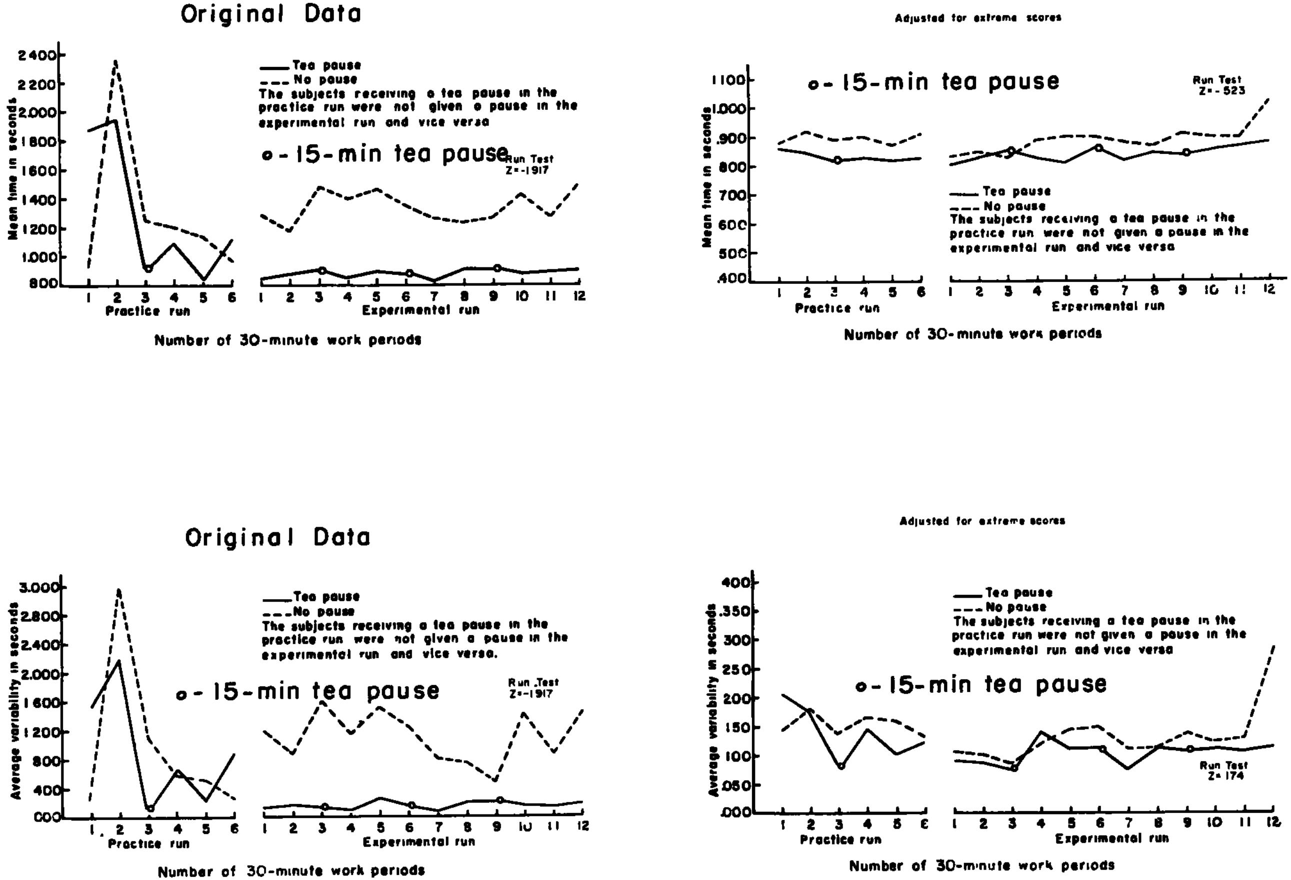


Figure 3. Steering efficiency vs time of driving.

Highly significant F-values were obtained for the 30-min intervals and the interaction between the treatment groups and the 30-min intervals for the steering score. Superior performance was demonstrated by the refreshment-pause group. This group was also superior in steering efficiency, as indicated by highly significant F-values for 30-min intervals and interaction.

A chi-square analysis of the number of extreme scores resulted in a value significant at the 5 percent level for stop light response time. The no-pause group had the greater number of extreme values.

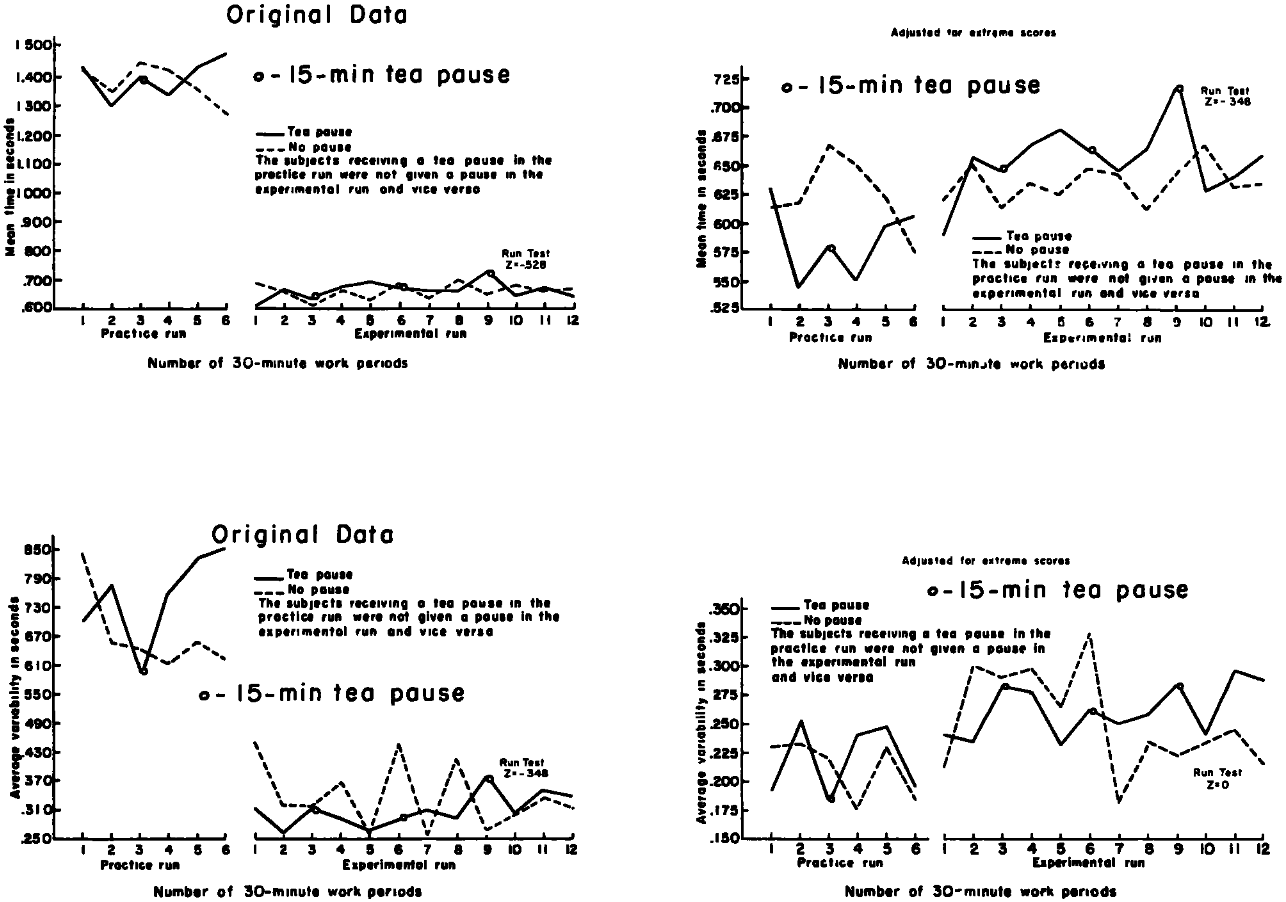
Figure 4.



Stop Light Response Time

(1) Mean time in seconds - original data
(2) Mean time in seconds - data adjusted for extreme scores
(3) Average variability in seconds - original data
(4) Average variability in seconds - data adjusted for extreme scores

Figure 5.



Train Reaction Time

(1) Mean time in seconds - original data
(2) Mean time in seconds - data adjusted
(3) Average variability in seconds - original data
(4) Average variability in seconds - data adjusted for extreme scores

TABLE 3

VARIABLES STUDIED DURING THE INTRANSIT PERIOD

Variable	F			Superior Group
	Treatment Groups	30-Min. Intervals	Inter-action	
Belt revolutions	2.319	27.490[2]	0.473	
Steering	3.318	29.326[2]	2.649[2]	RP[1]
Steering efficiency	1.229	12.627[2]	2.679[2]	RP
Stop light response time:				
Mean:				
Original data	1.747	0.488	0.384	
Adjusted for extreme scores	1.609	1.774	0.871	
Average variability:				
Original data	1.688	0.673	0.623	
Adjusted for extreme scores	2.527	1.250	0.667	
Train reaction time:				
Mean:				
Original data	0.004	0.483	0.862	
Adjusted for extreme scores	0.242	0.773	0.727	
Average variability:				
Original data	0.600	0.905	0.917	
Adjusted for extreme scores	0.207	1.417	1.333	

[1] Refresehment-pause group.
[2] Significant at the 1 percent level.

Each driver was given an opportunity to evaluate his own level of efficiency at the end of the experimental period. No group differences of sufficient magnitude to be statistically significant were revealed. Tiredness of eyes, physical discomfort, and nervousness were listed most frequently by the drivers as indicators of their own level of efficiency.

DISCUSSION OF RESULTS

It is in the area covered by the variables studied during the intransit period that the refreshment pause shows the most significant and potentially beneficial relationship to driving performance. It would seem that the refreshment pause not only increases the effectiveness of steering per half hour of simulated driving, but also tends to delay the onset and reduce the extent of decrement in performance. The same is true for the number of belt revolutions or simulated distance traveled each half hour of driving. This might tend to indicate higher concentration and greater zest on the part of the refreshment-pause group.

The subjects in the no-pause group made essentially the same estimate of their driving efficiency, as inferred from their responses to the subjective evaluation items obtained at the end of the 6-hr driving period, as that made by the subjects in the refreshment-pause group. No significant differences were shown. Because analyses of the objectively gathered data did reveal statistically significant differences between the two treatment groups with respect to several of the variables considered in this investigation, it appears that the individual driver is not always able to or at least does not always accurately estimate his own level of efficiency. A driver may have reduced efficiency and not be aware of it.

The time of onset of decrement differed among the variables studied. In several instances a decline appeared after 1 hr of driving. It was not until 2 hr of driving had elapsed that decrement was shown for all variables considered.

The results seem to indicate that it would be well for the average motorist to limit the amount of continuous driving to a period of from 1 to 2 hr. It would seem that to be conservative one should stop for rest every hour.

CONCLUSIONS

Within the limitations of the design, number of subjects, and other conditions of this study, the following tentative conclusions may be drawn:

1. The onset of work decrement occurs within the first 2 hr of simulated automobile driving task.
2. The effect of a refreshment pause can be detected most readily by measurements made while the subject is actually performing the task.
3. A refreshment pause substantially prolongs the onset of fatigue and reduces the extent of work decrement resulting from a prolonged period of simulated automobile driving performance task.
4. Drivers either cannot or do not accurately evaluate subjectively their own level of driving performance efficiency.
5. Variability in performance seems to be an indicator of the efficiency level of performance of the type studied.
6. A refresehment pause will increase maximum efficiency of performance.
7. The main hypothesis, that a periodic refreshment pause has no effect on the simulated automobile driving performance of automobile operators, can be rejected for three of the variables studied during the intransit period (steering, belt revolutions, and steering efficiency).
8. A driver's efficiency may be lowered without his being aware of it.
9. Drivers become less efficient after 2 hr of continuous driving.

ACKNOWLEDGMENT

The study reported herein was made possible through a grant for driving research from Thos. J. Lipton, Inc., and was made in the Driving Research Laboratory, Iowa State College, Ames, Iowa.

REFERENCES

1. "Too Long at the Wheel; A Study of Exhaustion and Drowsiness as They Affect Traffic Accidents." National Safety Council, Chicago (1935).
2. Ryan, A.H., and Warner, M., "The Effect of Automobile Driving on the Reactions of the Driver." Amer. Journ. of Psychol., 48:403-421 (July 1936).
3. Lauer, A.R., et al., "Selection Tests for Army Operators of Light and Heavy Motor Vehicles." U.S. Army, PRB Tech. Res. Rep. 1081.
4. Lauer, A.R., and Suhr, V.W., "Effect of Rest Pauses and Refreshment on Driving Efficiency." HRB Bull. 152 (1957).
5. Grubbs, F.E., "Sample Criteria for Testing Outlying Observations." Anna. Math. Stati., 21: 27-58 (Mar. 1950).