HIGHWAY RESEARCH RECORD

Number 366

Traffic Control and Driver Information

11 Reports

Subject Areas

51 Highway Safety
53 Traffic Control and Operations

HIGHWAY RESEARCH BOARD

DIVISION OF ENGINEERING NATIONAL RESEARCH COUNCIL
NATIONAL ACADEMY OF SCIENCES—NATIONAL ACADEMY OF ENGINEERING

WASHINGTON, D.C. 1971

ISBN 0-309-01981-8

Price: $4.00

Available from

Highway Research Board
National Academy of Sciences
2101 Constitution Avenue
Washington, D.C. 20418

CONTENTS

CRITERIA FOR NO-PASSING ZONES

G.W. Van Valkenburg and Harold L. Michael 1
Discussion
James L. Foley, Jr. 15
Edward F. Kearney 18
Authors' Closure 19

REAL-TIME NETWORK DECOMPOSITION
AND SUBNETWORK INTERFACING

R.J. Walinchus 20

PLANNING CONSIDERATIONS FOR TRAFFIC SIGNAL SYSTEMS

M.G. Davis and W.B. Hirsch 29

LEGIBILITY AND BRIGHTNESS IN SIGN DESIGN

Bernard Adler and Arthur L. Straub 37

DETERMINATION OF SIGN LETTER SIZE REQUIREMENTS FOR
NIGHT LEGIBILITY BY COMPUTER SIMULATION

Gerhart F. King 48

EVALUATION OF REAL-TIME VISUAL INFORMATION
DISPLAYS FOR URBAN FREEWAYS

Conrad L. Dudek and Hal B. Jones 64
Discussion
Marshall Jacks 75
F. Lehman 76
Authors' Closure 77

TRAFFIC DELAY AND WARRANTS FOR CONTROL DEVICES

Walter C. Vodrazka, Clyde E. Lee, and Herman E. Haenel 79

MEDIAN VISIBILITY IMPROVEMENTS: NEEDS,
METHODS, AND TRENDS

Jason C. Yu 92

DRIVER INFORMATION NEEDS

T.M. Allen, H. Lunenfeld, and G.J. Alexander 102

A COMPARISON OF SELECTED TRAFFIC INFORMATION DEVICES

Gerald C. Hoff 116

AN EVALUATION OF A REAR-MOUNTED VEHICLE
SPEED INDICATOR (Abridgment)

Clinton Jolliffe, Carl Graf, and David Alden 130

SPONSORSHIP OF THIS RECORD

Group 3—OPERATION AND MAINTENANCE OF TRANSPORTATION FACILITIES
Harold L. Michael, Purdue University, Lafayette, Indiana, chairman

Committee on Traffic Control Devices
Robert E. Connor, Federal Highway Administration, U.S. Department of Transportation, chairman
John L. Barker, O. I. Bermant, Robert L. Bleyl, Charles E. Dare, Robert N. Doble, J. Robert Doughty, John Elstad, Richard J. Folkers, Roy D. Fonda, Gordon R. Garner, Michael J. Gittens, Albert L. Godfrey, Sr., Alan T. Gonseth, Robert L. Gordon, Rudolph J. Israel, Barnard C. Johnson, Anthony C. Kanz, Ken F. Kobetsky, Leonard I. Lann, Robert D. McMillen, Joseph A. Mickes, J. P. Mills, Jr., Zoltan A. Nemeth, James I. Taylor, Francis E. Twiss, Robert E. Wert, Earl C. Williams, Jr., Robert M. Williston, Jason C. Yu

Committee on Motorist Information Systems
Gerson J. Alexander, Federal Highway Administration, U.S. Department of Transportation, chairman
Terrence M. Allen, Herbert J. Bauer, Wallace G. Berger, Robert E. Burns, John W. Eberhard, Eugene Farber, Donald A. Gordon, Ellis C. Henry, Gerald C. Hoff, Robert Hostetter, Larry L. Jenny, Arthur L. Korotkin, Ezra S. Krendel, Frederick G. Lehman, Harold Lunenfeld, Edward W. Melchen, Jr., Richard M. Michaels, Peter B. Moreland, Robert M. Nicholson, Robert C. O'Connell, Lawrence A. Powers, Arthur W. Roberts, J. E. Uhlaner, Donald L. Woods, Stuart Wright

K. B. Johns, Highway Research Board staff

The sponsoring committee is identified by a footnote on the first page of each report.

FOREWORD

Traffic engineers in particular, and operations specialists in general, will find the papers in this RECORD of interest and value in their day-to-day activities. The principal thrust of these studies is in how to make signals, signs, and pavement markings better serve the interests of improved traffic flow and safety. The findings can be very helpful in guiding practitioners to improved criteria for design and use in order to ensure greater safety and convenience to drivers.

In the opening paper, Van Valkenburg and Michael report the results of their research in the controversial area of no-passing zones, i.e., the short-zone versus the long-zone concept. State practices have generally been in accord with the short-zone concept, in which driving on the left side of the no-passing line is prohibited. The researchers conclude that the long-zone concept, which allows completion of passing maneuvers within the no-passing zone, should be adopted, and they developed criteria for marking the long zones and model legislation to allow implementation of the concept. Thoughtful discussions and the authors' closure further enhance the reader's understanding of this important work.

Walinchus presents a summary of new methods of decomposing or subdividing networks, and interfacing the resulting subnetworks, to improve real-time urban traffic control. Criteria for defining and selecting subnetworks and procedures for grouping them are discussed.

The use of computerized traffic signal systems is gaining in popularity, which led Davis and Hirsch to suggest that sound system engineering processes need to be applied in planning the systems to ensure successful operation. Application of the systems engineering process to signal systems is emphasized, and examples are given that relate to control system elements.

Adler and Straub concerned themselves with sign letter height considerations necessary to satisfy night legibility requirements. They concluded that required letter heights must exceed the 50-ft-per-in. rule in order to give good night legibility. They also feel that sign brightness can vary greatly because of a number of variables and that each sign should be treated individually as a separate design problem.

Using a computer simulation program, King analyzed 63 signs on a 20-mile Interstate highway for minimum required legibility under daylight and under high-beam and low-beam illumination. Letter size deficiencies were indicated for large numbers of these signs, even though they met prescribed requirements. Deficiencies were most pronounced for low-beam illumination and were especially severe for overhead signs, as might be expected.

By presenting questionnaires and slides to 505 motorists, Dudek and Jones determined their comprehension of and preferences for real-time visual information displays for urban freeways. Findings were that the motorists favored simplicity, use of color or other unique design features, and a design clearly distinguishable from other freeway signs. Discussions of the study techniques, as well as the study findings, extend the usefulness of the work.

Warrants for the installation of four-way stop control and validation of proposed warrants for installation of actuated signal control are presented by Vodrazka, Lee, and Haenel as results of their use of a unique digital delay data recorder. The device can record information from observers or from signal control devices in a form directly suitable for computer processing. Other applications for the delay recorder are discussed.

Starting with a nationwide survey, Yu studied median visibility considerations and current median delineation practices. He presents suggestions for improvements in median visibility, especially during nighttime.

Allen, Lunenfeld, and Alexander undertook an analysis of the task of motor vehicle driving to determine the structure of the many subtasks a driver performs. They found that these subtasks fall into a hierarchical scale from micro-performance levels such as steering and speed control to macro-performance levels such as trip planning and route following. The responses of drivers are varied as demanded by situations, including load-shedding of higher-level subtasks when required.

Motorists' preferences among six methods of receiving real-time freeway information were determined in a study by Hoff. Both visual and audible devices were used, and the researcher found preference for the visual. The subjects were also asked to make diversion decisions based on the information given, and the differences in their responses to the six methods were found to be small in a practical sense.

In the final paper, Jolliffe, Graf, and Alden report on their evaluation of a rear-mounted vehicle speed indicator in comparison with conventional rear lighting systems. The judgment of speed and distance by the driver of the following car was studied. Speed judgment errors were fewer with the speed indicator, but distance judgment errors showed no significant difference.

CRITERIA FOR NO-PASSING ZONES

G. W. Van Valkenburg, Parsons, Brinckerhoff, Quade and Douglas; and
Harold L. Michael, Purdue University

The concept currently used by most states for establishing and marking no-passing zones on two-lane highways legally prohibits motorists from driving on the left side of a yellow line throughout the length of a no-passing zone. The shortcomings of this concept, called the short-zone concept, are well known. It is physically impossible for motorists always to complete a passing maneuver without crossing the yellow line because of the limited visibility of no-passing zone signs and pavement markings. Furthermore, the crossing of a yellow line to complete a passing maneuver begun prior to the beginning of a no-passing zone is not an unsafe practice. An alternative to the short-zone concept is one that allows the yellow line to be crossed for the purpose of completing a passing maneuver. This concept, called the long-zone concept, prohibits the beginning of a passing maneuver in a marked no-passing zone. The purpose of this study was to determine which concept should be adopted to ensure maximum safety and comfort for the motoring public and to determine appropriate criteria and legislation to implement the recommended concept. The results of the research indicate that the long-zone concept, which legally allows the completion of a passing maneuver within a no-passing zone, should be adopted. Criteria for marking no-passing zones and a model law required to implement the concept were developed.

•DESPITE the current emphasis on building freeways, expressways, and superhighways, the bulk of the rural highway network throughout the United States is still the two-lane, two-way highway. At least 90 percent of the total rural mileage is two-lane, and much of this mileage was constructed before modern geometric design standards were established. Consequently the horizontal and vertical alignments create hazards that frequently are the indirect cause of many accidents.

A contributing factor to accidents occurring on two-lane, two-way highways is the limited sight distance, due to poor alignment, that exists on these roads. Sight distance is especially important on two-lane, two-way highways because the passing maneuver requires the use of the lane normally occupied by oncoming traffic. This constitutes a constant danger to the two-lane highway user.

To reduce this danger, traffic engineers for many years have established and marked no-passing zones with yellow paint and with "Do Not Pass" signs to warn drivers of impending sight restrictions. Laws regulating the behavior of motorists within these zones have been passed in every state to preserve the general welfare and safety of the motoring public.

Obviously, warnings of inadequate sight distances for passing on such highways should be clear, and motorists should always be certain of the meaning of such warnings. The criteria for establishment of no-passing zones and the exact meaning of such markings, however, are not uniform and can confuse motorists.

Many states have experimented with the use of additional marking devices to warn of impending no-passing zones. Perhaps the most popular device is the pennant-shaped "No-Passing Zone" sign mounted on the left side of the pavement. In 1967 three states

Sponsored by Committee on Traffic Control Devices and presented at the 50th Annual Meeting.

(Iowa, North Dakota, and South Dakota) were using this sign, and numerous other states were experimenting with it (11). Although the pennant-shaped sign is not in the 1961 Manual on Uniform Traffic Control Devices (MUTCD), it is included in the draft of a revised edition to be published in 1971.

Other devices that have been studied include a broken yellow line and semicircular blobs painted on the pavement preceding the solid yellow line. In Great Britain, large arrows have been painted on the pavement to direct traffic back to the proper lane (11).

The nature of the problem is apparent, but the solution has not been found. Usually studies have shown only a small reduction, if any, in the number of violations of the no-passing zone by the use of additional warning devices (11, 18). Traffic engineers have perhaps been addressing themselves to the wrong question. Instead of asking how to reduce or prevent violations of the no-passing zone, perhaps the question should be, Is it always dangerous to the motoring public when vehicles cross a yellow line? For example, is it dangerous to pass a farm tractor that is moving at a speed of 10 mph through a no-passing zone when it is obvious that there is ample distance free of obstructions or oncoming traffic in which to pass? Is it dangerous to finish a passing maneuver within a no-passing zone? Or, is it more dangerous to slam on the brakes when a no-passing zone is seen midway into a passing maneuver or to swerve abruptly in front of a passed vehicle to avoid crossing a yellow line?

Traffic laws that prohibit driving on the left side of an applicable yellow line throughout its length constitute what is known as the short-zone concept. An alternative to this is the long-zone concept, which prohibits the beginning of a passing maneuver within a no-passing zone.

The short-zone concept is contained within the recommended policy of the Uniform Vehicle Code (UVC) and the MUTCD. Consequently, most states have laws that incorporate the short-zone concept. Only a few states specifically allow the completion of a passing maneuver within a no-passing zone (10).

CRITERIA REVIEW

The 1961 edition of MUTCD contains criteria or warrants for the establishment of no-passing zones on two-lane and three-lane two-way highways. The criteria stipulate that, when the sight distance is less than a specified amount, a no-passing zone should be established (Fig. 1).

Changes in the MUTCD warrants were proposed in the early discussions for the new MUTCD (2). The proposals were not accepted, and the new MUTCD will contain the same minimum sight distances for no-passing zones as the 1961 edition. The sight distances are known to be inadequate for safe passing, however, and the problems associated with no-passing zones have made this topic a frequent matter of study and discussion by concerned committees of the American Association of State Highway Officials, the Highway Research Board, and the Institute of Traffic Engineers.

PURPOSE AND SCOPE OF PROJECT

The purpose of this research project was to improve the safety and efficiency of two-lane, two-way highways by improving the no-passing zone regulations and procedures. This involved two basic goals:

1. Determine the optimum warrants or criteria for the establishment of no-passing zones at horizontal and vertical curves on two-lane, two-way highways; and
2. Determine the necessary legislation to provide a legal and fair basis for the enforcement of restrictions on the passing maneuver, established according to these warrants.

METHODOLOGY

Passing Distance

Two distances are of primary importance in the determination of the sight distance needed to pass another vehicle: the distance traversed by the passing vehicle and the

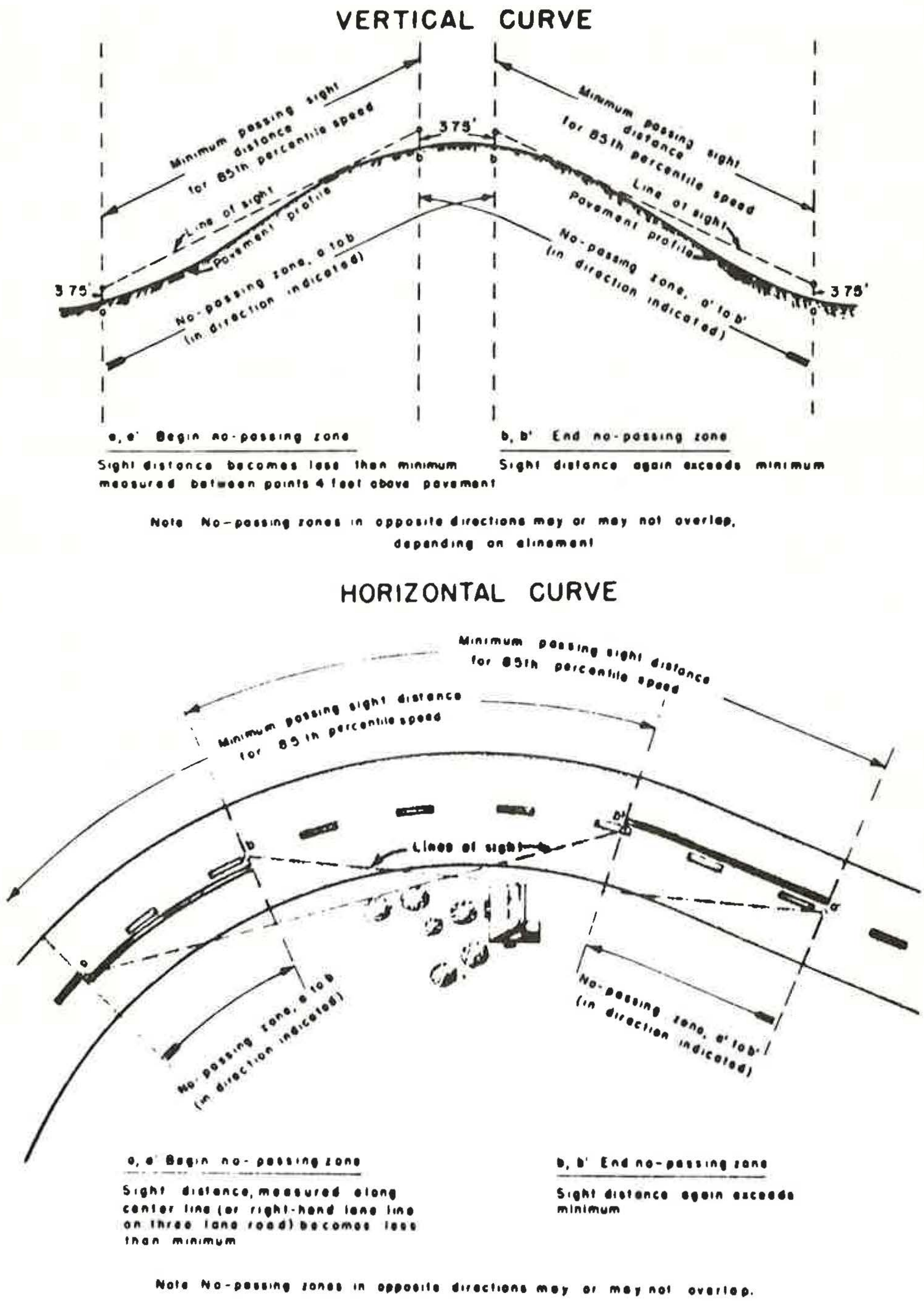


Figure 1. Determination of no-passing zones (Source: Manual on Uniform Traffic Control Devices).

distance traversed by an oncoming vehicle while the passing vehicle is in the "wrong" lane. This second distance is a function of the time needed to complete the passing maneuver, which is dependent on the speed and distance traversed by the passing vehicle.

The passing maneuver is shown in Figure 2. The first part of the passing maneuver, S_0, is the distance required to move abreast of the overtaken vehicle. The S_0 can be disregarded when calculating the minimum sight distance required for establishing no-passing zones. During this phase of the passing maneuver it is still possible to apply the brakes and pull back into the proper lane if an obstruction or oncoming vehicle comes into view. The exact location of this point (the so-called point of no return) may vary for each individual and among individuals depending on the characteristics of the passing vehicle and the speed of the passed vehicle and/or the speed of an approaching vehicle. However, it is generally assumed that the point of no return occurs when the passing vehicle is abreast or nearly abreast of the vehicle being passed. Based initially on personal judgment and subsequently confirmed through observation, the point chosen for this project occurs where the rear bumper of the passed vehicle is abreast of the middle of the passing vehicle. This is shown as point A in Figure 2. It was assumed that if a vehicle is at or beyond this point, the driver will determine generally that it is safer and easier to continue and complete the passing maneuver than to apply the brakes and pull back into position behind the vehicle being passed.

The minimum required sight distance to be determined by this research project was considered to be the sum of the following distances, as shown in Figure 2:

S_1 = The distance traveled by a passing vehicle between the point of no return and the point where it is completely clear of the "wrong" lane used by opposing traffic.

S_2 = The distance traversed by an oncoming vehicle while the passing car occupies the "wrong" lane as previously described.

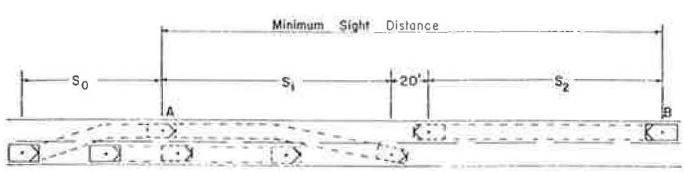
20 ft = An absolute minimum clearance distance between vehicles that would allow the two vehicles to avoid a head-on collision if the other assumptions were all met.

It was necessary to perform extensive field investigations of the passing maneuver to determine S_1 and S_2. The distance and time taken for passing maneuvers were observed by driving a test car at various speeds over selected sections of rural highways.

Test Roads

It was assumed that there may be a difference in the length and speed of passing maneuvers on different types of roads. Some of the features of a road that might introduce a bias include horizontal and vertical alignment, width and condition of pavement, the number and length of passing zones, and the volume and speed of traffic on the road.

Obviously, it was not feasible to test the effect of all possible variables. However, one important variable—the available sight distance conditions on a road—could be tested



ASSUMPTIONS:

1. The overtaken vehicle travels at a constant speed.
2. The oncoming vehicle reaches point B when the passing vehicle reaches point A.
3. The minimum sight distance is the sum of the distances $S_1 + 20' + S_2$

Figure 2. The passing maneuver.

if test roads of different geometric designs were chosen. For this reason three test roads, each 5 to 6 miles long and each having different visual restrictions, were chosen. Each test road had two test sections, one in each direction.

Test road S. R. 43N is a 5.53-mile long portion of State Route 43 located about 8 miles north of West Lafayette. The horizontal alignment on this stretch of road is generally straight with numerous vertical curves that restrict the sight distance on the southern end. There are five no-passing zones totaling 1.53 miles in the northbound direction and four no-passing zones totaling 1.40 miles in the southbound direction. About 28 percent of the road has a sight distance of less than 1,500 ft.

Test road S. R. 43S is a 6.20-mile portion of State Route 43 located about 7 miles south of Lafayette. There are five no-passing zones in each direction totaling 2.72 miles in the northbound direction and 2.82 miles in the southbound direction. About 40 percent of the road has a sight distance of less than 1,500 ft.

Test road S. R. 25 is a 5.4-mile portion of State Route 25 located northeast of Lafayette. The road has many hills and horizontal curves that restrict sight distance; 63 percent of the road has a sight distance of less than 1,500 ft. There are eight no-passing zones totaling 1.81 miles in the northbound direction and nine no-passing zones totaling 1.53 miles in the southbound direction.

Equipment and Personnel

The test car used throughout the experiment was a blue, 1962, 4-door Chevrolet sedan owned by Purdue University. A Stewart Warner survey speedometer with an odometer that reads to one-hundredth of a mile (52.8 ft) was mounted under the dashboard where it could be seen easily by both the driver and a passenger sitting in the front seat. A stopwatch was used to measure the time used during the passing maneuver. The same personnel, consisting of a driver and recorder, were used throughout the experiment.

Experimental Procedure

Numerous test runs were made by the test vehicle over the test roads to measure the lengths of the passing maneuvers and the time to complete a pass. The odometer was reset to zero at the beginning of each test run at the same beginning point for each test section. Therefore, the location of each passing maneuver within the test section could be plotted.

The type of vehicle and type of pass were noted for each pass. For instance, a pass made by a foreign car, pickup truck, single-unit truck, or semi-trailer truck was noted. It was noted also if the finish of a pass maneuver was hurried or forced by the presence of an oncoming vehicle or yellow line. Obviously, this was a judgment factor, but in most cases the abrupt, unnatural movement of the passing vehicle could be discerned easily.

Test runs were made only when the pavement was dry between the off-peak hours of 9:30 a. m. and 3:30 p. m. Monday through Saturday during the months of January, February, and March 1969.

The speed of the test car was maintained constant throughout each test run. Data were collected for three speedometer readings of the test car—40, 50, and 65 mph. The actual speeds of the test car corresponding to these speedometer readings were 38, 47, and 61 mph respectively. These speeds span the range of average traffic speeds that are usually found on two-lane, two-way roads during the off-peak hours.

The distance to pass was determined by taking a reading of the odometer when a vehicle was at the point of no return and taking another reading when the back wheels of the test car passed over the point where the left rear wheel of the passing vehicle crossed the centerline. The difference between these two readings gave a close approximation of the distance taken to pass.

The time to pass was determined by starting the stopwatch when the passing vehicle reached the point of no return and stopping it when the passing vehicle crossed the centerline as previously described. The decision of when the passing vehicle was at the beginning and ending points of the passing maneuver was made always by the driver of the test car. The driver also operated the stopwatch to minimize error due to perception and reaction time. The duty of the recorder was to read the odometer on the instruction of the driver and to record the readings.

A sight distance survey was made for each test section with sight distance measured along the centerline of the highway. For this survey the height of eye and target was 3.75 ft above the highway, in accordance with MUTCD criteria.

Speed studies were also made on the test sections to determine the speed distribution of traffic. The location in each case was on a tangent, level portion of the road where there was no restriction to the passing maneuver. This type of location was picked because this is where passing maneuvers occur most frequently.

EXPERIMENTAL RESULTS AND ANALYSIS OF DATA

General Observations

More than 3,000 miles were driven to collect data on the length and speed of the passing maneuver. Information on 915 passing maneuvers was recorded over a period of 3 months. The locations of no-passing zones, passing maneuvers, and sight distance were plotted for each of the six test sections. A portion of one of these test sections is shown in Figure 3.

There were frequent violations of the no-passing zones; i. e., the passing vehicle crossed an applicable yellow line at some point. There were 104 known violations (12 percent of all passes) of which 85 or 82 percent were passing vehicles returning to the right lane after the beginning point of a no-passing zone. Most of the remaining violations occurred when a vehicle initiated a passing maneuver prior to the end of a no-passing zone, especially when the passed vehicle was traveling at a slow speed or where the no-passing zone had been unduly extended.

It was also observed that traffic did not pass where sight distance was low, whether marked or unmarked. It appears that most drivers do not make a passing maneuver judgment only on the basis of the absence of an oncoming vehicle and the absence of a yellow line. If drivers cannot see what they consider to be a safe distance in front of them, they will not initiate a passing maneuver even though there may be no yellow line to warn them. Such a situation occurred most noticeably on test road S. R. 43N in the vicinity of station 1.4. In this area there is a horizontal curve that is not marked by a yellow line; yet not a single pass was completed at any speed in this area. The motorists did not think that they could see far enough to make a safe passing maneuver (maximum sight distance at one point is only 1,100 ft).

Data Classification

The types of passing vehicles were separated into four groups: automobiles, pickups, single-unit trucks, and semi-trailer trucks. The number of passing maneuvers completed by pickups, trucks, and semi-trailer trucks totaled 67, 24, and 27 respectively for all types of passing maneuvers and on all roads. A statistical analysis comparing the length and speed of passing maneuvers by these various vehicles was not undertaken because there were not enough observations to warrant conclusions. However, from inspection of the mean lengths and speeds of the passing maneuvers (Table 1) it is evident that criteria cannot be evolved for all types of vehicles without increasing the lengths of no-passing zones beyond that which would be reasonable or tolerable. Therefore, the statistical analysis was confined to passing maneuvers of automobiles only.

The types of passes were separated into four basic categories. An "accelerative pass" was a pass by a motorist who for one reason or another slowed down to the speed of the test car and followed behind the test car before initiating the passing maneuver. A "fly pass" was a pass by a motorist who did not slow down to the speed of the test car but passed the car "on the fly."

"Voluntary return" is a term used to describe the completion of a pass by a motorist when there was nothing forcing him to return to the right-hand lane. A "forced return" indicated that the motorist was forced to return to the right-hand lane by the presence of an approaching vehicle or the beginning of a no-passing zone.

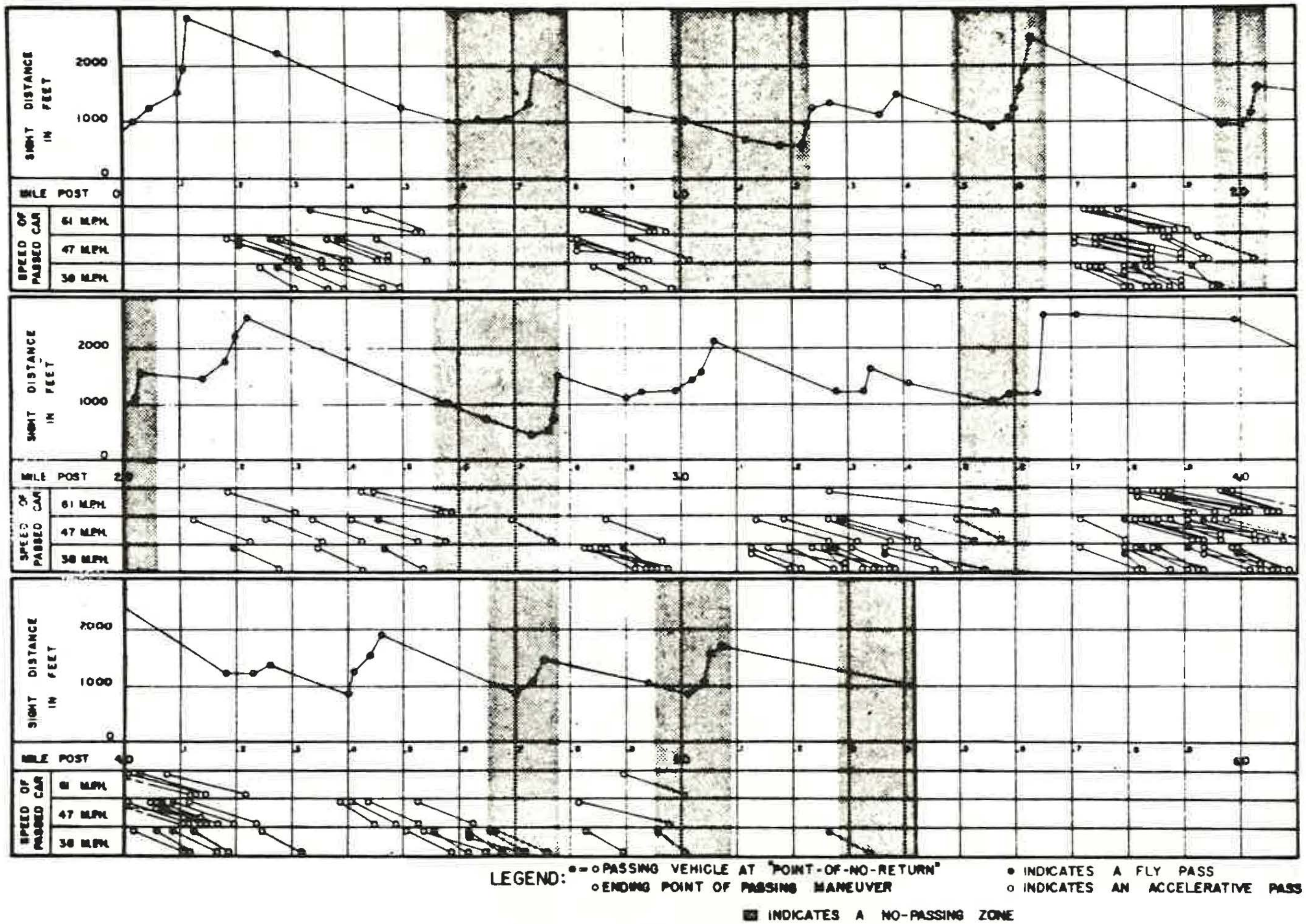


Figure 3. Location of passing maneuvers—S.R. 25 southbound.

TABLE 1

MEAN LENGTH AND SPEED OF PASSING MANEUVERS

Type of Pass	Speed of Passed Vehicle (mph)	Type of Passing Vehicle: Automobiles Dist. (feet)	Automobiles Speed (mph)	Pickup Trucks Dist. (feet)	Pickup Trucks Speed (mph)	Single-Unit Trucks Dist. (feet)	Single-Unit Trucks Speed (mph)	Trailer Trucks Dist. (feet)	Trailer Trucks Speed (mph)
Accelerative	38	496	49.6	531	48.1	666	48.2	906	42.7
Voluntary	47	618	56.9	693	54.6	642	52.0	965	52.2
Return	61	808	71.2	—	—	—	—	—	—
Flying	38	449	55.4	496	50.4	—	—	—	—
Voluntary	47	567	63.3	513	58.8	—	—	—	—
Return	61	619	74.4	—	—	—	—	—	—
Accelerative	38	339	49.1	—	—	—	—	—	—
Forced	47	430	61.3	—	—	—	—	—	—
Return	61	572	70.9	—	—	—	—	—	—
Flying	38	302	—	—	—	—	—	—	—
Forced	47	403	—	—	—	—	—	—	—
Return	61	—	—	—	—	—	—	—	—

Test Results

The mean length of passing maneuvers is given in Table 1 for the four types of passes: accelerative-voluntary return, flying-voluntary return, accelerative-forced return, and flying-forced return. Of these four types of passes, the mean length of the accelerative pass with a voluntary return by automobiles was consistently longer when passing the test car at speeds of 38, 47, and 61 mph than for the other types of passes. This is shown in Figure 4.

The mean speeds of passing vehicles of the various types for the four types of passing maneuvers are also given in Table 1. A plot of the mean speeds of the passing cars versus the speeds of the passed cars for three types of passing maneuvers is shown in Figure 5. From this it was apparent that the speed of passing vehicles in an accelerative type of pass with a voluntary return was lower than for other types of passes.

It was concluded, therefore, that both the speed and the length of an accelerative-voluntary return type of pass were most critical. Also, this type of pass occurred more

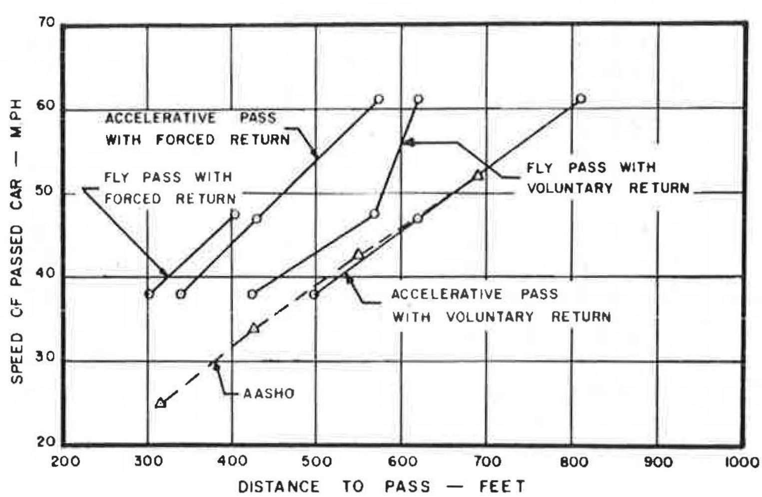


Figure 4. Length of passing maneuvers.

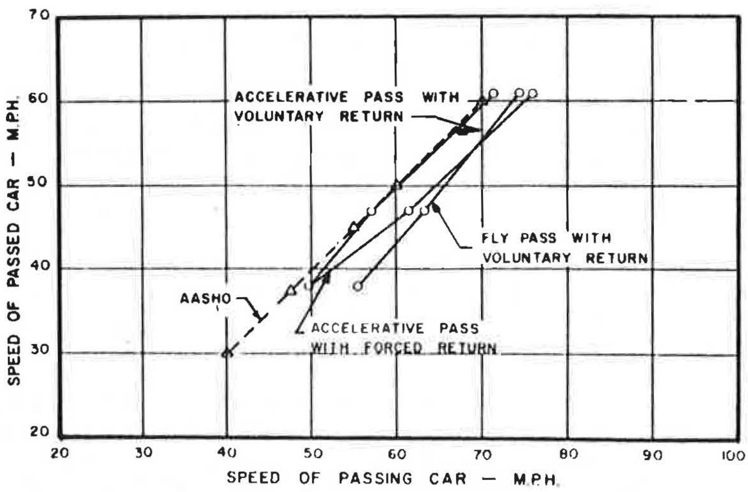


Figure 5. Speed of passing maneuvers.

frequently than any other. Therefore, the minimum sight distance requirements were based on the accelerative-voluntary return type of pass.

A comparison of the study results with AASHO criteria is shown in Figures 4 and 5. The dashed lines represent the AASHO criteria as taken from A Policy on Geometric Design of Rural Highways (4).

According to AASHO, "Speeds of overtaken vehicles were approximately 10 miles per hour less than speeds of passing vehicles." This was substantiated in this project and is shown in Figure 5. The dashed line in Figure 5 is a plot of the speed of the overtaken or passed car versus the speed of the passing car, assuming that the speed of the passing car is 10 mph faster than the passed car. As can be seen, the plot of the mean speed of accelerative-voluntary return type of pass nearly coincides with the AASHO plot.

Similar values taken from AASHO (4) were plotted by subtracting 10 mph from the average passing speed to obtain the average speed of the passed car. This plot, shown by a dashed line in Figure 4, falls very close to the plot of the mean length of the accelerative-voluntary type of pass obtained in this research project. The close proximity of these plots is coincidental. The AASHO plot is based on an acceleration pass with a forced or hurried return (4).

The AASHO data were also obtained from a study of selected no-passing zones by observing passing maneuvers at each from a fixed observation post (5). The procedure used in this research project made it possible to collect data under varying geometric conditions over test roads totaling about 17 miles in length. The AASHO data for the range of 60 to 70 mph for passing vehicles (corresponding to 50 to 60 mph for passed vehicles) is also based on extrapolated values. The results of this research, in addition to substantiating the accuracy of the AASHO data, suggest use of a different type of pass as the basis for no-passing criteria and extends the results to varying geometric conditions and higher passing speeds.

Statistical Analysis

The primary purpose of the statistical analyses was to determine if there was a significant difference in mean length to pass on various test roads and at various speeds. Through these analyses it was possible to determine what effects these variables had on the mean lengths and speeds and to place confidence limits on the test results.

An analysis of data within each test road concluded that overall it could be stated with a confidence level of 95 percent that there was no significant difference in the lengths of the accelerative-voluntary return type of passing maneuver in one direction over the other for a given test road. Test data taken in both directions were, therefore, combined.

Further analysis indicated that the individual test roads had an insignificant effect on the length of the passing maneuver. A maximum difference of mean passing distance between roads within the same speeds for the passed vehicle of only 0.015 mile or about 80 ft was found. On the other hand, the length of the passing maneuver increased significantly as the speed of the overtaken car increased, with a maximum difference of about 315 ft.

Throughout the study, it was the intent to be conservative. Passing maneuvers that were forced and subsequently much shorter—at least 150 ft (Fig. 4)—than those with a voluntary return were classified separately. On the other hand, passes by motorists who were obviously lazy in returning to the proper lane were included in the voluntary return classification.

It was the intent of this research project to develop criteria that could have a broad application to all roads. To do this, however, it would have been necessary to select a random sample of test roads throughout the United States. Therefore, the criteria, which were developed by combining data on all three test roads in this study, are theoretically applicable only to roads in the central area of Indiana. However, the statistical analyses indicate that the effect of roads on the length and speed of passing maneuvers is minimal. Therefore, it is suggested that the recommended criteria are sufficiently representative and conservative to be applicable to all roads.

Confidence limits on the mean length and speed of the passing maneuver were computed to provide an idea of how close the computed mean is to the true mean. One can be 95 percent confident that the true means of the length and speed of passing maneuvers are between the upper and lower limits given in Tables 2 and 3. The upper limit is the most important from a safety viewpoint. As can be seen in Table 2, the greater the speed of the overtaken car is, the greater is the variation in the length of the passing maneuver. The upper confidence limit at 61 mph for all roads combined was still only 0.007 mile (37 ft) longer than the mean length. From this it seems apparent that the test results are well within the accuracy necessary to establish safe criteria for no-passing zones.

Speed of Traffic on Test Roads

The speed studies showed that the mean speeds of traffic did not differ by more than 2 mph between test roads. The speed distribution curves indicated that about 70 percent of the traffic (15th to 85th percentile) traveled in a speed range of about 20 mph (48 to 68 mph). About 50 percent traveled within a range of ±5 mph of the mean speed of traffic.

TABLE 2

CONFIDENCE LIMITS OF THE MEAN DISTANCE TO PASS

Speed of Passed Car (mph)	Road	Number of Observations	Variance	Standard Deviation	Mean Distance (miles)	95 Percent Confidence Limits	
						Upper	Lower
61	S. R. 43N	41	0.00129	0.036	0.162	0.173	0.151
	S. R. 43S	40	0.00117	0.034	0.152	0.163	0.141
	S. R. 25	38	0.00173	0.042	0.146	0.157	0.135
	All	119	0.00141	0.038	0.153	0.160	0.147
47	S. R. 43N	79	0.00079	0.028	0.120	0.126	0.114
	S. R. 43S	61	0.00092	0.030	0.119	0.127	0.111
	S. R. 25	63	0.00077	0.028	0.112	0.119	0.105
	All	203	0.00083	0.029	0.117	0.121	0.113
38	S. R. 43N	58	0.00068	0.026	0.101	0.108	0.094
	S. R. 43S	60	0.00055	0.023	0.089	0.095	0.083
	S. R. 25	66	0.00049	0.022	0.092	0.097	0.087
	All	184	0.00058	0.024	0.094	0.098	0.090

TABLE 3

CONFIDENCE LIMITS OF THE MEAN SPEED OF THE PASSING VEHICLE

Speed of Passed Car (mph)	Road	Number of Observations	Variance	Standard Deviation	Mean Speed (mph)	95 Percent Confidence Limits	
						Upper	Lower
61	S. R. 43N	40	13.4276	3.6644	69.38	70.55	68.21
	S. R. 43S	39	17.7997	4.2190	70.78	72.14	69.42
	S. R. 25	35	13.6759	3.6981	71.63	72.90	70.36
	All	114	15.6016	3.9499	70.55	71.28	69.82
47	S. R. 43S	30	15.5175	3.9392	55.74	57.20	54.27
	S. R. 25	55	26.4066	5.1387	57.99	59.39	56.60
	All	85	23.5114	4.8489	57.20	58.24	56.15
38	S. R. 43S	31	26.8275	5.1795	48.60	50.50	46.70
	S. R. 25	52	22.2244	4.7143	50.57	51.89	49.26
	All	83	24.5611	4.9559	49.84	50.92	48.75

The frequency at which a vehicle will be passed is a function of its speed. Considerable passing of vehicles traveling less than the mean speed will likely occur while fewer vehicles traveling above the mean speed will be passed. Approximately 75 percent of the passing maneuvers, in fact, were noted to be of vehicles traveling at the mean speed or less. Therefore, it was decided to base no-passing zone criteria on the sight distance required to pass an automobile traveling at the mean speed of traffic. It must be assumed that drivers who pass a vehicle traveling above the average speed of traffic will realize the danger associated with this pass decision and will exercise appropriate safety precautions.

The speed of the oncoming vehicle (which may be out of sight) is an unknown quantity to the driver who is about to pass another vehicle. To base the minimum sight distance requirements on the average speed of oncoming vehicles might be dangerous because half of the approaching vehicles would be traveling faster than the average speed. Therefore, it seems logical to choose a speed that would include most of the oncoming traffic. Obviously, it is not practical to design for the 100th percentile speed. Therefore, it is simply a matter of judgment as to which speed to choose. The decision is not too critical, however, because the difference in speed between the 85th and 90th percentile, for instance, would be only about 2 mph.

Because the 85th percentile speed is often used in traffic engineering, this value was chosen for the speed of oncoming traffic in this study. The 85th percentile speed varied between 5 and 7 mph above the average speed on the test roads. This is also confirmed by annual speed studies conducted by Purdue University (1). Therefore, a speed of 7 mph faster than the average speed of the traffic was used as the speed of the oncoming vehicle.

Minimum Sight Distance

The minimum sight distance required to safely pass another vehicle is the sum of three distances, as follows: S_1, the distance to pass; S_2, the distance traveled by an oncoming car during that pass; and S_3, a clearance between the passing vehicle and the oncoming vehicle. The distance needed to pass and the speed of the passing vehicle have been established and are shown in Figures 4 and 5. Values were taken from these figures for each incremental speed, and the duration of the passing maneuvers could be calculated from the distance and speed of the passing maneuvers. From this the distance S_2 was calculated.

The total resulting minimum required sight distance is shown in Figure 6. The dashed line indicates extrapolated values outside the limits of this study.

Comparison With MUTCD Criteria

The sight distance criteria both from the 1961 edition of MUTCD and from an early draft of the proposed new MUTCD are shown in Figure 6. The MUTCD minimum sight

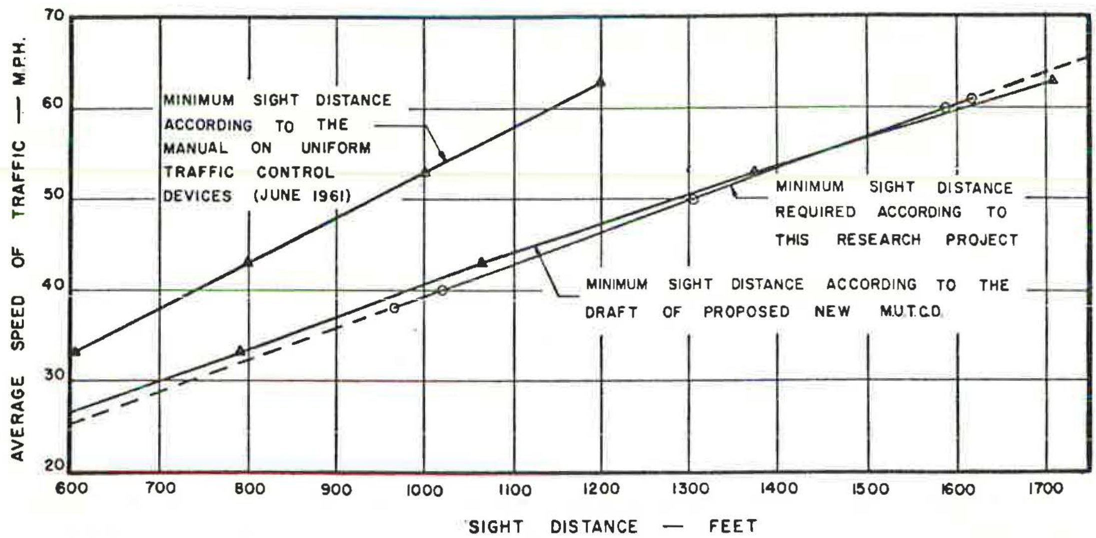


Figure 6. Minimum sight distance required to pass.

distances are stated for the 85th percentile speed of traffic, whereas the minimum sight distances developed in this research project are for average speed of traffic, i.e., the speed of the passed vehicle. As noted previously, the 85th percentile speed of traffic on two-lane, two-way state arterial highways in Indiana is approximately 7 mph faster than the average speed. As a consequence the minimum sight distances required by the MUTCD were plotted in Figure 6 with speeds 7 mph slower than the stated 85th percentile speeds for comparison with the average speeds used in this study. It is apparent that the proposed, but later rejected, MUTCD minimum sight distances coincide with the distances established in this research project.

The proposed new MUTCD draft also is associated with the same regulation as the 1961 MUTCD recommendations concerning the crossing of yellow lines. It is recommended in the MUTCD that an applicable yellow line not be crossed at any time. In effect this extends each no-passing zone by several hundred feet.

As an example, assume a motorist begins to pass a vehicle that is traveling at a speed of 60 mph and just as his vehicle reaches the point of no return he sees a no-passing zone sign ahead. At that moment he has the choice of braking the car to fall back into line or continuing the pass. Assuming that the pass is normal and average as defined in this study, he will need to be approximately 800 ft from the no-passing zone to be able to complete the pass and avoid crossing the yellow line (Fig. 4). Ordinarily a no-passing sign can only be seen about 300 to 400 ft away, and a yellow line is even less visible. The motorist would most likely be trapped into crossing the yellow line and would thereby become an offender of the law.

Many motorists are aware of the law, and rather than continue a normal passing maneuver they swerve abruptly in front of the passed car to avoid crossing the yellow line. This dangerous movement was observed frequently during this experiment. It was obvious that such a maneuver did not contribute to the safety and pleasure of either the passed or passing motorists and their passengers.

Minimum Distance Between No-Passing Zones

The minimum distance between no-passing zones that should be allowed without making one continuous zone is stipulated in the 1961 MUTCD as 400 ft. The early proposals for the new MUTCD would have increased this distance, especially at higher speeds, in line with requirements of the short-zone concept. If this minimum distance were increased, the effect would be to increase further the length of no-passing zones and decrease the legal opportunities to pass slow-moving vehicles. Consequently, capacity would be reduced and the frustration of motorists following slow-moving vehicles would be increased.

The distance required to initiate a passing maneuver was investigated. Assuming that the initial phase of the passing maneuver is equal to one-third of the total distance to pass (as assumed by AASHO), one-half of the distance S_1 as measured in this study would correspond to the length of the initial phase. This distance represents the average distance that a motorist would need to accelerate and arrive at the point of no return if he were watching and waiting for the end of the no-passing zone to appear. These distances were found to vary from 190 ft at 30 mph to 460 ft at 70 mph. It appears that the existing 400-ft minimum distance is adequate and could even be reduced for slower speeds under the long-zone concept.

CONCLUSIONS AND RECOMMENDATIONS

The most obvious conclusion reached during this research project was the inadequacy of the short-zone concept currently utilized by nearly all of the states. The large number of motorists who illegally cross an applicable yellow line should not all be classed as law offenders. The law is clearly inconsistent with the physical capabilities of the driver and vehicle. Consequently the law cannot always be obeyed. Such a situation can only contribute to disregard of laws in general and utter frustration for the unfortunate few who are apprehended.

The long-zone concept allows the completion of a passing maneuver across the yellow line. If the motorist is so far into the maneuver that a severe braking action is required to stop the maneuver in order to avoid crossing the barrier line, the motorist is allowed to continue the maneuver, for by design such as continuation would be safe. The beginning of a no-passing zone becoming visible during a passing maneuver would, however, provide a cautionary warning similar to the yellow caution light in traffic signals. The approach of an applicable no-passing zone during a passing maneuver should demand safe and reasonable action on the part of the motorist. Enforcement against violators requires no more judgment on the part of law enforcement personnel than the enforcement of traffic signal regulations.

There is another important aspect to the problem that cannot be ignored. Uniformity of traffic laws and criteria throughout the nation is a necessary and desirable goal. It is true that several years will be required before all states would or could change their laws to adopt the long-zone no-passing concept. However, the shortcomings of the short-zone concept are well known, and many individuals will not be convinced that their state should adopt a law that is known to be unsatisfactory. But most important, many motorists either are unable or do not want to comply with the short-zone concept, as evidenced by the large number of violations of no-passing zones in this study and others (11, 18).

The logical alternative is to allow the applicable yellow line to be crossed for the purpose of finishing a passing maneuver that was well under way before the beginning of a no-passing zone was reached. This can be achieved through the universal adoption of laws and criteria to implement the long-zone concept. Criteria and legislation that might be adopted to implement the findings of this research are given in the following. The major changes in the wording of the Manual on Uniform Traffic Control Devices and Uniform Vehicle Code suggested are italicized (except the tabulated values which are also changes).

Criteria for No-Passing Zones at Curves (MUTCD)

> . . . Where centerlines are installed, a curve warrants a no-passing zone and should be so marked where the sight distance is equal to or less than that listed below for the prevailing (off-peak) *average* speed:

Average Speed, Off-Peak (mph)	Minimum Passing Sight Distance (ft)
30 and under	750
31 to 35	900
36 to 40	1,050
41 to 45	1,200
46 to 50	1,300
51 to 55	1,450
56 to 60	1,600
61 to 65	1,750
66 to 70	1,900

The following table indicates the minimum distance between no-passing zone markings necessary for initiation of a passing maneuver:

Average Speed, Off-Peak (mph)	Minimum Distance Between Zones (ft)
30 and under	250
31 to 40	300
41 to 50	350
51 to 60	400
61 to 70	450

Where these minimum distances cannot be provided, the no-passing zone markings should be connected to form one continuous zone.

Legislation (UVC)

The following change in the Uniform Vehicle Code, Section 11-307, No-Passing Zones, is suggested so that the long-zone concept may be incorporated into no-passing zone legislation:

Model Law—No-Passing Zones

(a) (No change from current wording.)

(b) Where signs or markings are in place to define a no-passing zone as set forth in paragraph (a) no driver shall at any time drive on the left side of the roadway within such no-passing zone or on the left side of any pavement striping designed to mark such no-passing zone *except for the purpose of safely completing a passing maneuver begun prior to the beginning point of such a zone.*

(c) (No change from current wording.)

REFERENCES

1. Hejal, S. S. Traffic Speed Report No. 86. Joint Highway Research Project, Purdue University, 1968.
2. Manual on Uniform Traffic Control Devices for Streets and Highways. National Joint Committee on Uniform Traffic Control Devices, U. S. Department of Commerce, 1961, pp. 122-127.
3. Part III of Draft of New MUTCD, unpublished. National Joint Committee on Uniform Traffic Control Devices, 1968, pp. 11-15.
4. A Policy on Geometric Design of Rural Highways. American Association of State Highway Officials, 1965, pp. 140-152.

5. Prisk, C. W. Passing Practices on Rural Highways. HRB Proc., Vol. 21, 1941, pp. 366-378.
6. Review of Driver Eye Height as Related to Registered Passenger Vehicles. Traffic Research Section, Michigan Department of State Highways, unpublished, 1967.
7. Uniform Vehicle Code. National Committee on Uniform Traffic Laws and Ordinances, 1968.
8. Indiana Motor Vehicle Laws. Central Publishing Co., Indianapolis, 1965.
9. Indiana Motor Vehicle Laws, Revised. Central Publishing Co., Indianapolis, 1967.
10. Uniform Vehicle Code: Rules of the Road With Statutory Annotations. National Committee on Uniform Traffic Laws and Ordinances, 1967, pp. 192-234.
11. Kelly, Don C., and Sidnell, John E. D. The Researchable Aspects of No-Passing Zone Signing and Marking. Department of Civil Engineering, University of Kentucky, 1967.
12. Marshall, F. C. Determining and Marking No-Passing Zones. Public Works, Vol. 92, No. 2, 1966, pp. 76-78.
13. A Plan and Procedure for Locating No-Passing Zones. Traffic Engineering Division, Arizona Highway Department, 1962.
14. Bartels, W. J. No-Passing Zone Procedures. Traffic Engineering, Vol. 28, No. 7, 1958, pp. 15-16.
15. A Review of the Criteria for Marking No-Passing Zones. Traffic Division, Michigan Department of State Highways, 1963.
16. A Review of the Criteria for Marking No-Passing Zones—1968. Traffic and Safety Division, Michigan Department of State Highways, unpublished, 1968.
17. Ostle, Bernard. Statistics in Research. Iowa State University Press, Ames, 1963.
18. Study of No-Passing Zone Signing. Traffic Research Section, Michigan Department of State Highways, 1965.
19. Van Valkenburg, G. W. No-Passing Zones: Criteria, Legislation and Location. Joint Highway Research Project, Research Report No. 19, Purdue University, June 1969.

DISCUSSION

James L. Foley, Jr., Federal Highway Administration

This report provides much-needed research for reevaluating pavement-marking law and practices. As Figure 6 shows, the authors' recommended minimum passing sight distances closely match those originally proposed for the 1970 edition of the Manual on Uniform Traffic Control Devices. However, because supporting research was not available at the time, the proposal was not accepted by the National Joint Committee. This study produced needed up-to-date data on current passing practices of drivers.

In discussing the concepts for marking no-passing barriers, the authors describe the short-zone concept as being based on the legal precept that driving to the left of any portion of the no-passing barrier is illegal. They describe the long-zone concept as being based on the legal precept that the passing maneuver legally may be completed by crossing the yellow barrier. A third concept also merits evaluation—that the no-passing barriers be advisory only and be used only as guides for enforcement purposes. In this case, the passing prohibition would be enforceable only at those locations where the proper regulatory signs have been installed. The logic for this concept is that pavement markings, because they are subject to wear and are hard to see during inclement weather, are not suitable devices for enforcement purposes. The standard "Do Not Pass" regulatory sign would be required for this purpose.

Perhaps the item that should be examined more closely, in regard to making a change from the present short-zone concept, is the degree of hazard in initiating a passing maneuver in those areas just in advance of the beginning of the no-passing barrier.

The authors indicate that we have a problem in the manner of marking no-passing zones. I concur with their conclusions that the long-zone concept is a safer legal base on which to design the no-passing zones. It is interesting to note that the Uniform Vehicle Code position was adopted in 1956. Since then the vehicle and, to a degree, the highway have undergone significant changes. An indication of the seriousness of the passing problem is the fact that between 18 and 21 percent of the fatalities on rural highways involve two vehicles traveling in opposite directions. Use of these values should not be misinterpreted. We cannot presume that they were all passing accidents. Included in this percentage are all head-on crashes, including those in which one vehicle crosses the centerline for reasons other than to pass. It also includes wrong-way driving on divided highways. Fortunately, this percentage has decreased in recent years. A part of this is undoubtedly due to the increasing share of the traffic using the Interstate and other limited-access facilities. On the other hand, the increasing use of the Interstate System requires that the guidance we give the motorist off the Interstate System be the best possible because fewer and fewer drivers have learned to cope with the passing situation on two-lane highways.

If motorists had accurate depth perception and could estimate the speed of an approaching vehicle, perhaps only a minimum-length sight distance marking would be required to permit a driver to make a safe passing maneuver. Farber and Silver (20) indicate that drivers are fairly competent at judging the distance element but are very poor at determining approaching vehicle speed. This becomes critical when the passing driver has begun his maneuver and the oncoming vehicle suddenly appears in sight.

Two new ideas have been proposed by the authors. The first is the development of a valuable new term, point of no return, to describe a long-held concept. The other is the use of the average speed of traffic as the speed of the passed vehicle and as the basis on which no-passing zones should be marked. I subscribe to the point-of-no-return concept and believe that most drivers operate on that assumption. However, there is some question whether the point of no return is truly the point at which slowing to fall back behind the passed vehicle is more difficult or more hazardous than completing the pass. It is entirely possible that the total closing distance traveled by opposing vehicles while the passing vehicle is braking from the point of no return to fall back behind the passed vehicle would be less than the distance traveled by both vehicles while completing the passing maneuver. This, of course, would be dependent on (a) the passed vehicle maintaining its speed and not slowing to assist the passing vehicle in completing his maneuver, (b) the acceleration characteristics of the passing vehicle, and (c) the actual speed of the passed vehicle. To be valid, this comparison should be based on a forced return rather than a voluntary return. Research similar to that used by the authors for this paper would be helpful to provide the data needed for this comparison.

When forced to return to the right-hand lane in a limited distance because of the pavement marking, either drivers will make the return regardless of the dangerous "cutting-in" process, or they may ignore the legal regulation and enter the no-passing zone to the left of the no-passing barrier marking. The more desirable alternative would certainly be the latter, even though it tends to diminish a driver's respect for the barrier. The use of the barrier line concept as recommended by the authors (permitting the completion of the passing maneuver over the barrier line) would appear to reduce the attention the driver must give to markings. This would permit his concentration on the possible appearance of an opposing vehicle on the road ahead. The 12 percent violation of no-passing barriers suggests that many drivers are relying on their judgment in passing rather than on the absence or existence of a barrier line. The driver who endangers himself or others by trying to comply with the legal statute by "cutting in" probably creates a greater hazard than the one who violates the short-zone concept.

According to the averages developed by the authors, the traveled distance required for a forced return is about 30 percent shorter than for a voluntary return (Table 1). Inasmuch as safety is the paramount consideration in marking no-passing zones, we should look at the potential hazard of the present marking practices and compare it with that proposed by the authors. Because of the high correlation between the authors' findings and the data developed by AASHO for design of two-lane highways, AASHO definitions can be used for this comparison. The AASHO term d_4 (the distance traveled by the opposing vehicle) is considered to be two-thirds of d_2 and equal to the distance traveled by the passing vehicle while completing its maneuver. The sum of these two values relates very closely to the minimum sight distance +20 ft as used by the authors ($S_1 + S_2$ + 20 ft). Figure 5 compares these values with the values used in the 1961 Manual on Uniform Traffic Control Devices and the authors' recommendations. It appears that the current practice (1961 Manual) does not provide an adequate safety factor for completing the pass. On the other hand, the proposal of the authors would provide a comfortable margin of safety. Figure 6 compares the present practice with $^7/_{10}$ of the values for the voluntary return; i. e., 0.7 ($d_4 + {}^2/_3\, d_2$). Examination of Table 1 indicates that passing distances with a forced return approximate $^7/_{10}$ of the passing distances with voluntary returns. This seems to indicate why the current practice is not completely unworkable. Thus, it would appear that the current practice assumes a forced return.

Another item that needs further investigation is the distance between the passing vehicle and the approaching vehicle at which the passing driver begins to feel uncomfortable and forces a return to the right-hand lane. In Figure 2, the 20-ft distance between driver positions selected as an absolute minimum between the passing and the approaching vehicles may be a little short. It is considerably less than the distance in the AASHO Policy on Geometric Design for Rural Highways, and no indication was given of how this distance was selected. For design, the AASHO policy uses a clearance distance varying according to the speed of the two approaching vehicles, ranging from 100 ft at 30 mph to 250 ft at 53 mph. In a study performed in Saskatchewan (21), a distance of 40 ft was used for the clearance. From my viewpoint, even 40 ft appears less than most drivers would tolerate and tends to result in erratic maneuvers.

In terrain where no-passing zones are frequent, the use of the short-zone concept requires the inclusion of criteria for connecting adjacent zones when time to initiate and complete a pass does not exist. The authors have recommended values ($^1/_2\, S_1$) for minimum distances between no-passing zones. I agree that, under the concept that permits completion of a passing maneuver once initiated, the use of much shorter distances between successive no-passing zones is appropriate. Because the driver is aware that the end of the zone is being approached, the use of $^1/_3\, d_2$, or perhaps even a shorter value, would allow safe completion of the passing maneuver.

This sampling of the passing practices of drivers improves our understanding of the problem. What is still unknown is the extent to which drivers place themselves in danger when performing a passing maneuver. If we are to attempt to place a barrier line at a given location on the basis of a single speed representing approaching traffic and a single speed for the traffic being passed, it undoubtedly will not fit all situations and just as obviously will not be suitable for precise interpretation by enforcement personnel. The recommendation of the authors proposing use of median speed for calculating the length of no-passing zones provides a desirable factor of safety for marking existing highways. For newly constructed two-lane highways, markings must be based on assumed speeds. Perhaps the design speed should be assumed to be the 85th percentile, and the markings should be based on a median speed 7 mph slower. As an example, for initial markings, use an assumed median speed equal to 7 mph slower than the design speed of the new highway. As a compromise between the short-zone, rigid-enforcement concept and the advisory concept, the authors' proposal seems appropriate.

This research project provides a nucleus and, as such, is a valuable contribution to the engineering field. The authors are to be commended for developing a relatively simple research procedure. It would be desirable for others in different parts of the country to use the technique to add to the data base and to increase the usefulness of the recommendations.

References

20. Farber, E., and Silver, C. A. Knowledge of Oncoming Car Speed as Determiner of Driver's Passing Behavior. Highway Research Record 195, 1967, pp. 52-65.
21. Crawford, J., and Pollock, D. Passing Sight Distance Study for Pavement Markings. Saskatchewan Department of Highways.

Edward F. Kearney, National Committee on Uniform Traffic Laws and Ordinances

One of the principal conclusions of the paper is that our general driving rules should be changed to allow a motorist to complete a passing maneuver after he has entered an area designated by signs or markings as a no-passing zone. In my opinion, the information presented in the paper does not support that conclusion. Indeed, the existing rule requiring passing drivers to return to the right side prior to reaching the zone is the one more in accord with the authors' research findings.

Specifically, the paper indicates that of 915 passing maneuvers observed at three locations, drivers crossed the yellow line in only 12 percent of the cases. If we assume that this 12 percent refers to crossing the line at the conclusion of a pass upon entering a zone, it suggests that 88 percent of the passes were completed without using space on the left side of the roadway within a no-passing zone. This does not support statements appearing at several points in the paper suggesting that there were frequent violations of the no-passing zones, that a high percentage of motorists crossed the yellow line, and that many motorists either were unable or did not want to comply with the "short-zone" concept as evidenced by the large number of violations.

In 1968, the National Committee on Uniform Traffic Laws and Ordinances rejected, by a vote of two to one, a proposal to allow completion of a pass within a no-passing zone. One of the principal reasons for this rejection was that the requirement to return to the right side prior to reaching the zone is both well understood and complied with. The paper's revelation that at least 88 percent of the drivers passed without improper use of the left side of the roadway within a no-passing zone is very reassuring and completely justifies the National Committee's decision. Also the 12 percent reported in this study as not negotiating a timely return to the right side may be high because "passes by motorists who were obviously lazy in returning to the proper lane were included" and "passing maneuvers that were forced" may have been excluded and because the paper suggests that the zone markings may have been unduly extended.

At this point it may be well to consider the purpose of a no-passing zone. I think its purpose is to indicate an area on a two-lane, two-way roadway where driving on the left during any portion of a passing maneuver is so highly fraught with peril it must be prohibited. Any other statement of its purpose does not make sense because other rules of the road require driving on the right and prohibit any form of passing that would interfere with the safe operation of any overtaken vehicle or any vehicle approaching from the opposite direction.

The suggestion that passes can be safely completed within a no-passing zone indicates a lack of appreciation as to why such areas exist. Further, the suggestion that safe passes be allowed to be completed within a zone is meaningless because it is already required that all passes be accomplished without interfering with oncoming and overtaken vehicles. In addition, not only does the paper advocate allowing completion of a pass begun well before the commencement of a zone, the draft of the proposed model law would allow driving on the left side of a zone to complete a passing maneuver begun a few feet before the zone. Another curious feature of the draft is that it would not allow completion of a pass in an unmarked zone (such as within 100 ft of an intersection) but would allow such completion if the zone were marked.

Allowing a driver to complete a pass within a no-passing zone is a contradiction in terms. Although some people may think that "passing" includes only the act of actually passing another car, courts hold that "passing" includes moving left, overtaking, passing, or returning to the right side for the purpose of passing a moving vehicle. Because the paper suggests banning only the first of these four elements within a no-passing zone, it would allow the other three—hence the legal contradiction.

It seems to me that advocates of the "long-zone" theory of designating unsafe passing areas think that extending their length is necessary or desirable for advance notice of the existence of the zone. I disagree with this contention because such zones do not exist in isolation. They generally exist because of a hill, curve, intersection, or some other condition that makes passing unsafe. Usually, advance notice of these conditions is provided by appropriate warning signs. Also, signs placed on the left and right sides of the highway can convey sufficient advance notice of the zone without extending its length. One also should question the practicality of giving advance notice by paint on a level roadway.

Clearly, a driver passing to the left of another car on a two-way roadway has a heavy burden to discharge because he may not interfere with the safe operation of a vehicle approaching from the opposite direction nor the one being overtaken or passed. Safe and efficient highway travel demands that a passing driver return to the right side of the roadway prior to reaching the zone. The paper does not justify changing these rules or telling 110 million drivers that they may lawfully complete a pass once inside a no-passing zone. The better rule of the road is and should remain: Return to the right side prior to reaching the zone.

AUTHORS' CLOSURE

We thank each of the discussers for his thoughful comments. Special appreciation is due Mr. Foley for his complimentary remarks about our research and findings. We heartily agree that additional research, especially concerning accidents involving vehicles that initiate a pass just prior to a no-passing zone, would be advisable.

One of the discussers suggests that the short-zone regulation is well-obeyed by motorists and that the 12 percent violation statistic is proof of this. The fact is, however, that most passing maneuvers cannot be in violation of the law because of the locations on the road where passing maneuvers generally occur, i. e., where oncoming no-passing zones do not exist. Further, a very high percentage of those maneuvers that occur where there is a chance to be in violation are in violation. We cannot agree, therefore, that the data substantiate good observance of the present no-passing laws.

The same discusser also claims that to be to the left of a no-passing line at any time is "highly fraught with peril." This is not true for the greatest part of every passing situation. A vehicle that is completing a passing maneuver will require only about 100 ft to return safely to the right lane, not the total distance indicated as not available by the no-passing line. Contrary to what the short-zone law states, the no-passing line does not indicate that it is unsafe to be to the left of the line. It means only one thing: At every point along a given line a specific sight distance is not available, and therefore, in the normal passing situation, if you start a passing maneuver at this point you do not have adequate sight distance. The no-passing line does not mean that if you are halfway through, three-quarters of the way through, or nine-tenths of the way through the pass you cannot safely complete the maneuver. It is only applicable to the beginning of the passing maneuver. The completion of a pass initiated even only a few feet prior to a no-passing zone is a safe maneuver; it does not interfere with oncoming or overtaken vehicles if the no-passing zones are correctly marked. (They are not correctly marked under the short-zone concept.) On the other hand, a vehicle in the left lane that is swerved abruptly to the right to avoid violating a no-passing zone does interfere with the overtaken vehicle.

The most serious shortcoming of the present short-zone concept, however, is that at certain hazardous zones there is no marking at all. For example, on a 60-mph average-speed road it is clear that a minimum sight distance of 1,600 ft is required for safe passing. The short-zone distances given in MUTCD, however, provide for marking no-passing zones only if 1,150 ft are not available. Therefore, for this 60-mph road all locations where sight distances between 1,150 and 1,600 ft are available will not be marked under the short-zone concept, even though it is clear that minimum passing sight distance is not available. Such situations are often more common than very short sight distances on two-lane highways, and many unsafe passing areas are not marked under the short-zone concept. Under the long-zone concept they would be marked.

We subscribe to the suggestion that we should consider making no-passing zones advisory. It is ridiculous to tell a driver that it is unsafe to pass a slow-moving farm wagon where 1,000 ft of sight distance is available, and it will be treated in that way by many drivers.

It is possible that a change in the no-passing laws may require wording different from that suggested in order to be legally correct; however, such revisions could easily be made.

REAL-TIME NETWORK DECOMPOSITION AND SUBNETWORK INTERFACING

R. J. Walinchus, TRW Systems Group, Houston

New methods of decomposing networks and interfacing the resulting subnetworks are being developed as part of the continuing research and development effort to improve real-time urban traffic control. The basic concepts and alternatives are introduced in this paper. Network decomposition may be required when the network geometry is not uniform and/or when the traffic characteristics within an area are not uniform. Decomposition is not successful unless two tasks are accomplished: (a) subnetwork determination according to certain criteria and (b) interfacing the various subnetworks. The two general ways in which subnetworks may be defined are fixed definition and real-time subnetwork definition. Fixed subnetworks are defined from geometrical and other considerations. Real-time subnetwork definition is based on traffic dynamics. To define successfully subnetworks in real time, two elements must be present: (a) a criterion by which it can be determined that an intersection is a "candidate" for consideration as part of a separate subnetwork and (b) a procedure by which these candidate intersections can be grouped into a workable subnetwork. The proposed criterion for candidacy is offset error—the difference between a realizable and an idealized offset. Among the grouping procedures are two genuinely real-time methods (rectangular subset and connectible set) and one pseudo real-time method. Interfacing subnetworks can be a difficult task. In the special case where the two areas have the same signal cycle, interfacing is accomplished via a "delta offset." This delta offset is a volume-weighted average of the offset changes desired across the interface. With unequal cycles, a matching and resynchronization technique might be employed, or interfacing can be accomplished through a transition zone where the signals are traffic-actuated rather than operated on any specific cycle.

•THE urban traffic control system (UTCS) of the 1970s will be burdened with increasing demands for more effective traffic control. The Federal Highway Administration (FHWA), as part of its effort to improve urban traffic conditions, is establishing the UTCS Laboratory in Washington, D.C. The UTCS Laboratory will employ a digital computer, other off-the-shelf hardware, and existing stored-pattern software; however, it will also serve as a tool for the development of advanced control techniques.

The FHWA requested proposals (1) for the development of second-generation traffic-control software that could be operational in the Washington Laboratory in 1972. The prospectus recognized the need for subnetwork configurations. This need was identified by the Traffic Systems Office of TRW Systems Group under an earlier FHWA contract (2) and has been pursued by the author through in-house studies. Developments in real-time network decomposition and subnetwork interfacing techniques are introduced in this paper.

Sponsored by Committee on Traffic Control Devices and presented at the 50th Annual Meeting.

Real-time network control via first-generation techniques requires choosing an appropriate stored pattern of signal settings; second-generation network optimization requires detailed computation of the signal settings based on traffic dynamics. An example of a network optimization program is SIGOP (3), but SIGOP is an off-line, steady-state program as opposed to a real-time one.

Most network optimization programs contain an iteration routine or a matrix inversion, or both (e.g., subprogram OPTIMIZ in SIGOP). The computation time required for these operations is nonlinear with respect to the number of variables involved; i.e., doubling the number of variables will more than double the computational time required (quite often, the time increase is nearly exponential). Partitioning the network is a major benefit because solving several small problems will, in general, take less time than solving one large problem. Thus, partitioning of a network into subnetworks (fixed and/or real-time definition) makes optimum use of available time on the digital computer, which is used to periodically compute optimum signal settings.

The analytical basis of network optimization via subnetworks can be likened to manipulating a very large matrix to solve a big problem: It is easier and faster to partition the matrix (network) into submatrices (subnetworks). This partitioning may be required when the network geometry and/or the traffic characteristics are not uniform or when the optimization problem is too large to solve in real time. In addition, partitioning must be judicious; network decomposition and subnetwork interfacing must be done with care so that the optimization of individual subnetworks yield satisfactory results when integrated.

The decomposition of a general urban network into subnetworks will become an integral part of real-time urban traffic control. Network decomposition is not successful unless the following two tasks can be accomplished:

1. Subnetwork determination according to certain criteria, and
2. Interfacing the various subnetworks.

These two tasks are discussed in the following sections.

SUBNETWORK DETERMINATION

Two general ways in which a network may be decomposed into subnetworks are (a) fixed subnetwork definition and (b) real-time subnetwork definition. Fixed subnetworks arise from geometrical and other considerations. Real-time subnetwork definition is based on traffic dynamics.

For a general urban area, a combination of the two methods may be necessary because of the limited instrumentation that may be available and because of certain peculiarities in geometry that may exist in the area.

Fixed Subnetworks

There are a number of criteria that are used to establish fixed subnetworks. Typical decomposition criteria are (a) freeway access/service road; (b) major arterial into urban grid; (c) area of closely spaced intersections; (d) signal-independent flow area; (e) geometrical/political subdivisions; and (f) established traffic patterns. A few comments about each of these will illustrate their applicability for any given urban area.

Where urban streets empty onto freeways or merge into freeway service roads, a subnetwork boundary usually can be defined. Where a major arterial feeds into an urban grid, the arterial and the grid may be considered as different subnetworks with appropriate boundary conditions.

Certain areas are a maze of closely spaced signals. Often, these "dense" areas are surrounded by relatively long blocks with reasonably laminar flow. Decomposition around the middle of the long surrounding blocks is possible.

Certain areas and streets have sufficient feeders and turning movements so that traffic flow is rather consistent and virtually independent of the settings of surrounding signals. (In some cases, this condition is brought about by lack of instrumentation and signal control.) This is an ideal place to perform network decomposition. This criterion is usually related to the following one.

In many cities a river, creek, park, or institution provides for easy decomposition. In other areas, the political subdivisions between communities or counties may introduce mandatory subnetwork boundaries.

Finally, there may be areas where the pattern of behavior of traffic is "established" in some sense. The city traffic engineer may decide on subnetworks because "area A always requires a signal cycle of about X seconds and area B always needs one of about Y seconds." This is one of the less desirable criteria for subnetwork definition because it assumes constancy and will limit flexibility to optimize flow if conditions change.

Regardless of which criteria are used, the result is the same—fixed subnetworks. Some degree of flexibility can be obtained by storing in the computer several network "decomposition maps," any one of which can be called in by an evaluation of traffic conditions, by keyboard entry, by time of day, etc. This is discussed in more detail under the heading "Pseudo Real-Time Definition."

A more meaningful criterion for the definition of subnetworks is traffic dynamics. This implies real-time subnetwork definition, as will be discussed.

Real-Time Subnetwork Definition

In order to define successfully subnetworks in real-time, two elements must be present:

1. A criterion by which it can be determined that an intersection is a candidate for consideration as part of a separate subnetwork, and
2. A procedure by which these candidate intersections can be grouped into a workable subnetwork.

Criteria for Candidacy—Real-time subnetwork definition based on traffic conditions requires establishment of criteria by which an intersection can be judged to be a candidate for separation from the network proper.

One criterion for subnetwork candidacy is "offset error" at the intersection. This parameter is defined as the optimum offset when the intersection is considered as part of the whole network minus the idealized offset of the intersection with respect to adjacent intersections. (The "idealized offset" as used in SIGOP is a constant; in real-time urban traffic control it is dynamic and is based on speeds, queues, etc.) When the offset error is too large (compared to a constant or dynamic threshold value), this implies that the network solution is constraining the intersection to be inefficient on a local basis. This intersection is then a candidate for consideration as part of a separate subnetwork.

A variation of the preceding criterion is "weighted offset error," where the offset error on each approach to the intersection is weighted by the pertinent volume (rather, vehicles requiring service). In this way, an intersection with few vehicles requiring service need not be overemphasized because the offset error may be counterbalanced by spare green time in which the improperly phased traffic can be processed.

Other criteria that might be used to determine candidacy are intersection saturation level, traffic density, etc.

The criterion chosen depends to some extent on the amount of instrumentation and the accuracy of data for a given intersection. It is difficult to determine the candidacy of an intersection for which traffic dynamics data do not exist.

In the discussion that follows, assume that an appropriate criterion or combination of criteria has been chosen; the candidate intersections must be grouped into a workable subnetwork using one of the methods discussed.

Figure 1 will be useful in the following discussions. It shows a pattern of candidate (dark dot) intersections and a street-intersection numbering technique, namely via i and j. The north-south streets are specified by a specific value of i, and the east-west streets by j. An intersection is specified by an i-j pair, (i,j). Among all the intersections in the network, the candidate intersections form a set $\{(i^*,j^*)\}$. In Figure 1, this set of candidate intersections contains

$$\{(i^*,j^*)\} = \left\{\begin{matrix}(3,3), & (3,4), & (4,3), & (4,4), & (4,5) \\ (5,2), & (5,4), & (5,5), & (6,3), & (6,4)\end{matrix}\right\}$$

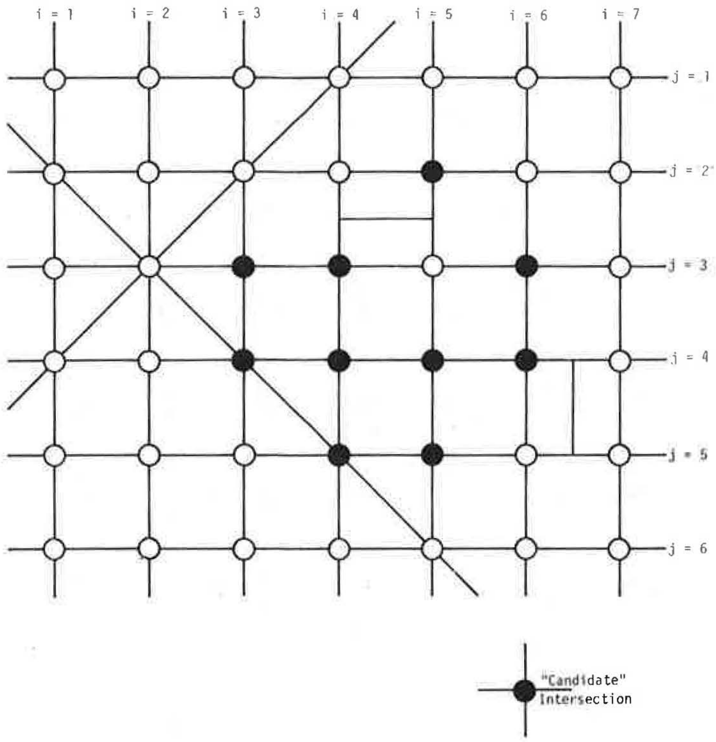


Figure 1. Intersections that are candidates for grouping into a subnetwork.

This street-intersection numbering technique and the set of candidate intersections are particularly useful in the "rectangular subset" method of real-time subnetwork definition.

Grouping Methods—Experimentation with three methods of grouping candidate intersections into workable subnetworks has been pursued. The methods are pseudo real-time, real-time (rectangular subset), and real-time (connectible set). The first two are being programmed and checked out.

Pseudo real-time subnetwork definition makes use of stored decomposition maps; the map that best fits current conditions is used. The two real-time methods, rectangular subset and connectible set, do what their names imply. In the first, a rectangular area containing all (or the highest density) of the candidate intersections is used to define a subnetwork. The second merely searches for the largest grouping of candidate intersections that are connected by optimizable links. More details on all three methods are given in the following.

Pseudo Real-Time Definition—This method uses one of the several stored "decomposition maps" that tell how the network should be partitioned into subnetworks. One version employs an index cross-referencing procedure to sort traffic data according to subnetworks. If a library of these maps is available, it becomes a matter of picking the best map to match current traffic conditions. If the stored maps handle the frequently encountered situations, matching is relatively easy. For example, consider Figure 2.

In the figure, the area within the dashed circle has given trouble in the past, and therefore the corresponding decomposition map has been stored in the library. The current traffic conditions show a high density of candidate (dark dot) intersections in this area and none outside. Therefore, the indicated decomposition map is the one best matched to current conditions, and a subnetwork is defined by the intersections within the dashed circle.

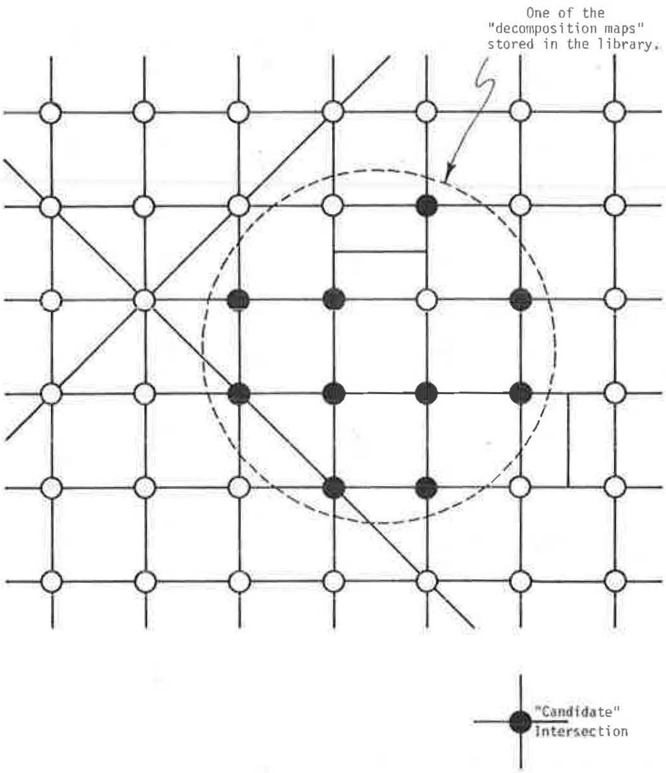


Figure 2. Pseudo real-time subnetwork definition: decomposition map matching.

The name pseudo real-time has been attached to this method because of the possibility of poor map matching (due to variable traffic conditions and/or an incomplete map library). The real-time methods define their own subnetwork boundaries.

Real-Time Definition: Rectangular Subset—This method searches for a rectangular area that includes all (or the highest density) of the candidate intersections.

By using the street-intersection numbering technique shown in Figure 1, it is relatively easy to determine the rectangular area enclosing all of the candidate intersections. All that need be done is to find the maximum and minimum i^* and j^* in the set $\{(i^*,j^*)\}$. Define the following quantities:

$$I_1 = \max_{\{(i^*,j^*)\}} i^* \quad , \quad I_2 = \min_{\{(i^*,j^*)\}} i^*$$

$$J_1 = \max_{\{(i^*,j^*)\}} j^* \quad , \quad J_2 = \min_{\{(i^*,j^*)\}} j^*$$

In Figure 1, these quantities have the values

$$I_1 = 6 \quad , \quad I_2 = 3$$

$$J_1 = 5 \quad , \quad J_2 = 2$$

The corresponding rectangular subnetwork that includes all of the candidate intersections is within the dashed area shown in Figure 3. This technique has been programmed.

A variation of the preceding technique is to find a smaller rectangular area (within the one above) that contains a higher density of candidate intersections. This is ac-

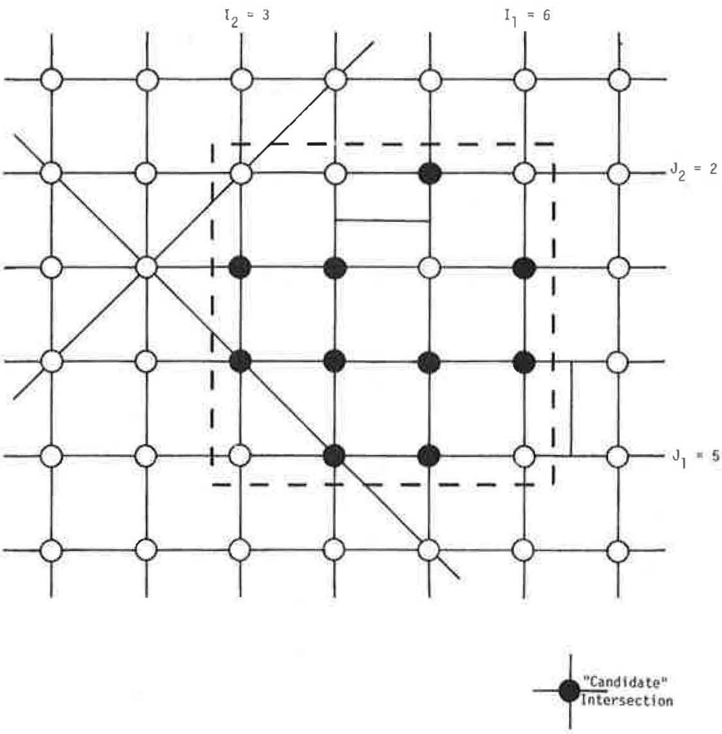


Figure 3. Subnetwork definition via rectangular subset containing all candidates.

complished simply (from a programming standpoint) by eliminating one street at a time from the sides of the full rectangular area; i.e., test

$$
\begin{aligned}
I_1' &= I_1 - 1 \text{ with } I_2, J_1, J_2 \text{ fixed} \\
I_2' &= I_2 + 1 \text{ with } I_1, J_1, J_2 \text{ fixed} \\
J_1' &= J_1 - 1 \text{ with } I_1, I_2, J_2 \text{ fixed} \\
J_2' &= J_2 + 1 \text{ with } I_1, I_2, J_1 \text{ fixed}
\end{aligned}
$$

and eliminate the one street that leaves the highest density of candidate intersections in the remaining rectangle. (The procedure can be applied, if necessary, more than once, provided the resulting rectangle does not yield a trivial or undesirable case.) Applying the procedure once to the situation in Figure 3 gives the situation shown in Figure 4 where

$$
I_1 = 6 \quad , \quad I_2 = 3
$$
$$
J_1 = 5 \quad , \quad J_2' = 3
$$

The "density" of candidate intersections has been increased from 0.625 (in Figure 3) to 0.75 (in Figure 4) while dropping only one candidate intersection. (That one intersection will be no worse off, but the three adjacent noncandidates will benefit by remaining part of the other area.) The corresponding smaller rectangular subnetwork is within the dashed area shown in Figure 4.

Real-Time Definition: Connectible Set—This method is an extension of the higher density concept; it is the most logical from a grouping standpoint, but not necessarily the best when considering the overall optimization problem.

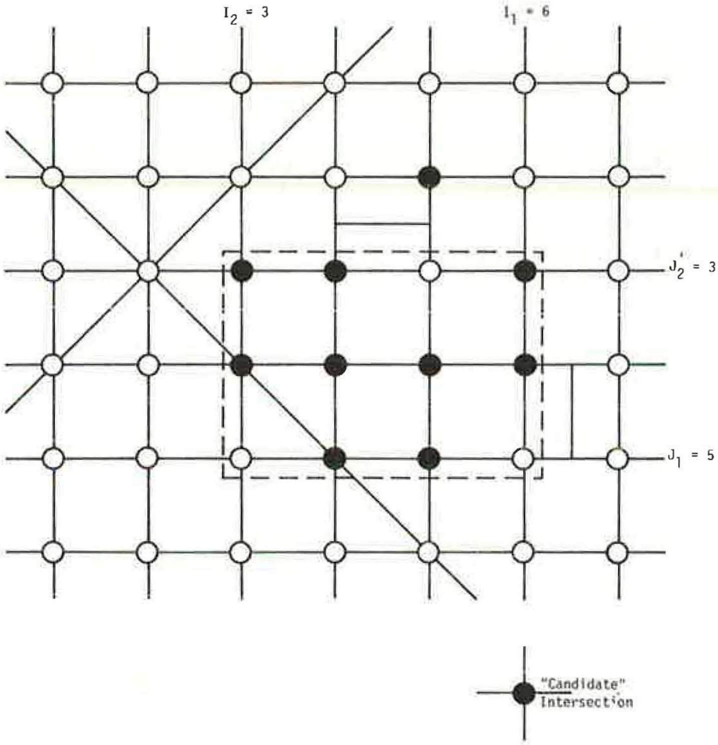


Figure 4. Rectangular subset containing higher density of candidates.

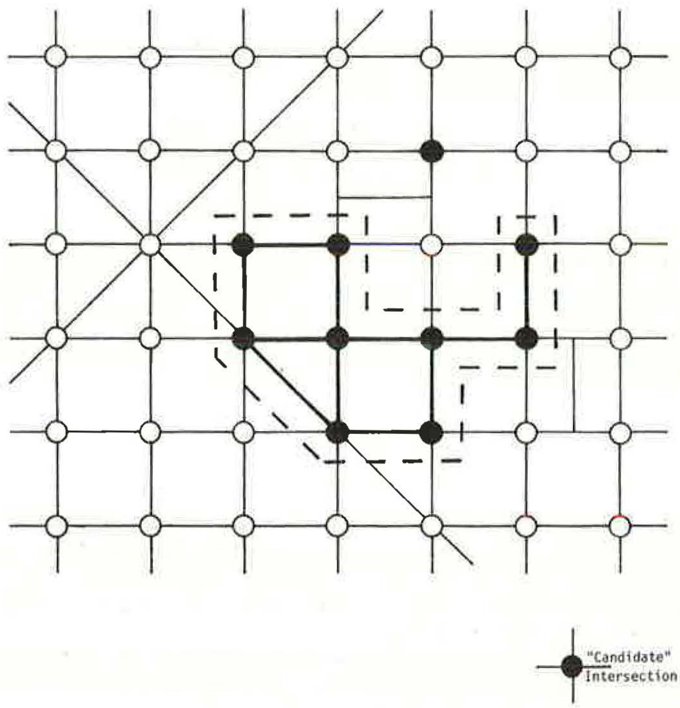


Figure 5. Largest connectible set of candidates.

It requires sorting through the list of candidate intersections and associated links to find the largest connectible set. (In this context, a pair of candidate intersections is "connectible" if there is an optimizable link and no other optimizable intersection between the pair.) Figure 5 shows the results of the method where the pertinent connecting/optimizable links are shown as dark lines. Note that one candidate intersection is not connected to the rest of the group.

The method is largely a sorting problem. The efficiency of the bookkeeping operations may depend a great deal on the numbering scheme for identifying intersections and links.

A drawback of the method is the irregular subnetwork boundary that may result (see dashed line in Figure 5). This will affect the efficiency of merging traffic smoothly at the boundary interface.

INTERFACING SUBNETWORKS

The preceding section dealt with the problem of defining subnetworks. Once the subnetworks are defined, traffic data can be sorted according to subnetwork, and each subnetwork can be optimized. Then comes the problem of interfacing subnetworks with one another and the rest of the network.

If a subnetwork and surrounding area have the same optimum signal cycle, the interfacing task is much simpler. With different cycles there are two approaches that can be taken for interfacing:

1. Initially match the offsets of the two areas for good traffic flow between them; then, as the two signal cycles move too far out of synchronism, disrupt the subnetwork cycle to force resynchronization; and
2. Introduce a "transition zone," at least one intersection wide between the two areas, letting the transition intersections operate almost on a traffic-actuated basis.

Interfacing Under Equal Cycles

When two areas have the same signal cycle, interfacing across their common boundary is relatively simple. It is a matter of introducing a "delta offset" to all the offsets within the subnetwork (or to the subnetwork master signal) to put its main-street green-on times in synchronization with the surrounding area.

The delta offset is computed via the equation

$$\text{Delta offset} = \frac{\displaystyle\sum_{\substack{\text{interfacing}\\ \text{links}}} (\text{Volume})\,(\text{Offset change desired})}{\displaystyle\sum_{\substack{\text{interfacing}\\ \text{links}}} (\text{Volume})}$$

where

Offset change desired = Offset of upstream intersection + Idealized offset between intersections based on free-flow travel time - Offset of downstream intersection, modulo the signal cycle

for each interfacing/boundary link. Thus, the delta offset is simply the volume-weighted average of the offset changes desired across the interface.

Encountering two adjacent areas with the same (or nearly the same) signal cycle introduces some interesting questions:

1. If the two areas have the same signal cycle, should they be combined into and optimized as one area?
2. If the two signal cycles are "nearly" the same, should (a) the two areas be combined into one, or (b) the two areas be kept separate but constrain the two signal cycles to be the same?

The traffic engineer will have to answer these questions before development of the subnetwork definition and interfacing program is completed.

For areas with different signal cycles, one of the two following approaches can be employed.

Interfacing by Matching/Resynchronization

In the matching/resynchronization technique, an initial matching of traffic flow is accomplished using the delta offset previously discussed. However, because the two signal cycles are different, the signal settings move out of phase so that traffic flow between the two areas is not synchronized. It is desirable to resynchronize the settings periodically if it is not too disruptive to traffic flow.

Periodic resynchronization could be performed as new optima are computed. A routine can be developed whereby subnetwork signal phases are appropriately extended or shortened to resynchronize the settings.

This technique might be required in cases where the subnetwork boundary/interface must be so abrupt as to prohibit introduction of a transition zone between areas.

Interfacing Through a Transition Zone

The other interfacing technique calls for a transition zone (one or two intersections deep) to be established between the subnetwork and the surrounding area. The signals in this zone do not operate on any specific cycle.

The phases of the transition-zone intersections are set on the basis of traffic demand. The demand can be determined by actuation (if instrumentation exists) or prediction (of platoons leaving the subnetwork or surrounding area and proceeding toward the transition zone). The prediction method of determining demand probably has to be used because of the limited amount of instrumentation that may be available in a given urban area, and because the instrumentation is not likely to be in any desired transition zone.

CONCLUSIONS

This paper has presented the basic concepts and alternatives for the real-time decomposition of networks into appropriate subnetworks and subsequent interfacing of the subnetworks. Some of these developments will be applied in the Washington laboratory beginning in 1972; others may be in use in cities by the mid-1970s.

ACKNOWLEDGMENTS

The author is principal investigator under a research and development contract between the Federal Highway Administration and the TRW Systems Group. The opinions, findings, and conclusions expressed are those of the author and not necessarily those of the Federal Highway Administration.

REFERENCES

1. Traffic Adaptive Network Signal Timing Program. U.S. Dept. of Transportation, Federal Highway Administration, RFP-121 (24-82), March 1970.
2. Final Report: System Analysis Methodology in Urban Traffic Control Systems. TRW Report No. 11644-HO14-RO-00, 30 June 1969. Available from National Technical Information Service, Springfield, Virginia, as PB 185 422.
3. SIGOP: Traffic Signal Optimization Program. Traffic Research Corp., New York, September 1966. Available from National Technical Information Center as PB 173 738.

PLANNING CONSIDERATIONS FOR TRAFFIC SIGNAL SYSTEMS

M. G. Davis and W. B. Hirsch,
Wilbur Smith and Associates, New York

Computerized traffic signal systems are receiving increasing consideration as a means of alleviating roadway congestion. Many of the systems that are being considered are based on the use of a digital computer. The increased scope and complexity of these signal systems require the application of sound system engineering processes if successful system operation is to be achieved. A summary of the systems engineering process is related, and its application to traffic signal systems is emphasized. Examples are given that examine several traffic control system elements with respect to configuration, reliability, and accuracy.

•IN response to the continual increase in urban traffic congestion, an expanding number of cities are investigating the use of computerized traffic signal systems to achieve optimum use of their existing road networks. Impetus toward the optimal use of existing road systems provided by the federally aided TOPICS program has also resulted in increased interest in computerized traffic control systems. Although the state of system control hardware far exceeds the present state of traffic control technology, justification of a real-time control system using a digital computer can often be reached through consideration of the following factors:

1. For a signal system totaling fifty intersections or more, the cost of a digital computer system is not greatly in excess of the cost of more common system configurations and may be less for equal control capability;
2. The existing state of traffic control technology is not reflected in conventional hardware, and the control algorithms implemented by conventional hardware are not easily altered by equipment modification; and
3. Because a traffic control system represents a substantial investment that should have a useful life of 10 to 20 years, the flexibility and computing power of the digital computer provide the control hadrware that is best able to incorporate advances in control theory.

Implementation of computer-oriented traffic control systems of the scope under consideration can best be achieved by the application of sound systems engineering technology. The systems approach includes the total and continual planning of the system from inception to phaseout. The need for this type of planning and constant evaluation lies in the complexity of the engineering task. The justification for this discipline is derived from the benefits to be gained by successful system implementation and the penalties to be suffered by unsuccessful implementation.

The computer-oriented traffic signal systems currently being planned and instituted differ in several important aspects from the systems of the past two decades. One area of significant difference is the increased size and scope of the equipment serving the new systems and its capacity for expansion and modification. A second difference of great importance is the potential ability of the system to perform optimum control

Sponsored by Committee on Traffic Control Devices and presented at the 50th Annual Meeting.

strategies based on modern control theory. Because the degree to which any control system can achieve optimization depends to a large extent on the accuracy and refinement with which the system can measure and evaluate the quantities to be optimized, the accuracy requirements and quantity of detector data can be expected to increase substantially. The increased data communications requirements, number of control and detection elements, severe environment in which control and detection elements must operate, and economic constraints create system problems that must be resolved to achieve reliable system operation.

The advanced state of hardware technology has added a new element of choice to the design of the traffic control system. Historically, the decisions in traffic control systems centered on whether the system was to be completely pretimed or was to be traffic responsive. Once this basic decision was made, the remaining decisions dealt largely with the selection of a manufacturer to supply the equipment. Because the functional differences among the various systems offered by the traffic equipment manufacturers tended to be negligible, the choice of the system supplier had only minimal effect on system planning, system operation, and system cost.

Today, the range of choices available is extremely wide and will probably grow. Previously the system control functions were integral with the hardware and were fixed; now the element of system software has added a completely new dimension to system design. The implications of this added element are extensive and, at times, subtle. Basically, the operation and efficiency of the entire control system is a function of the software. However, before software requirements can be defined, the overall requirements of the system must be determined. Because the choice of possible system requirements and functions has become quite broad and because all elements of the system can be affected by these basic requirements, it follows that a systematic procedure is required to achieve an optimum configuration of hardware and software.

Because software will perform a major role in the implementation of system requirements, the specifying of software becomes a key element in the system specifications. Historically, specifications for traffic signal systems have generally described the control functions of existing hardware or have treated the subject of how control was to be achieved superficially. Because there was little difference among the control algorithms available, these specifications may have been satisfactory. However, the complexity of digital computer control demands that precise specifications be made if the intent of the system requirements is to be met.

If the full potential of the computerized traffic system is to be realized, a continual program of control evaluation and modification must be established. The system requirements should necessarily reflect this need in both the hardware and the software specifications. In summary, the modern computer-based traffic control system can no longer be viewed as a piece of hardware selected from a catalog; it must be designed to meet the growing transportation problem by implementing the expanding traffic control technology. The complexities of the technology do not permit success to occur without adequate planning.

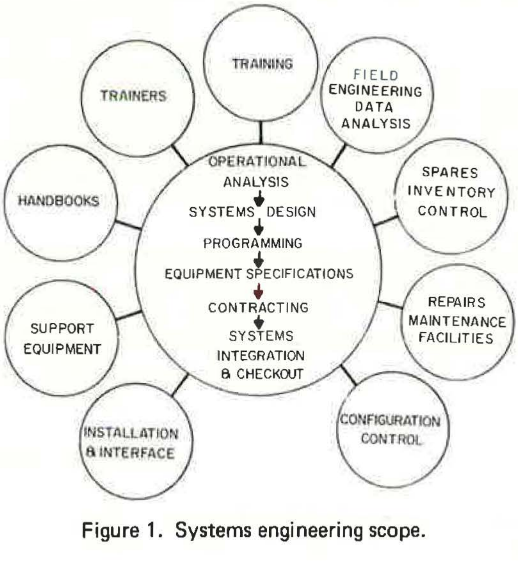


Figure 1. Systems engineering scope.

A system denotes the total resources brought to bear on a problem. These resources include both personnel and hardware and extend beyond equipment performance to include training, maintenance, operation, and evaluation procedures. Figure 1 shows the total involvement of the system engineering discipline. Involvement with the total problem is of itself not the key to successful system engineering. The benefits from systems engineering

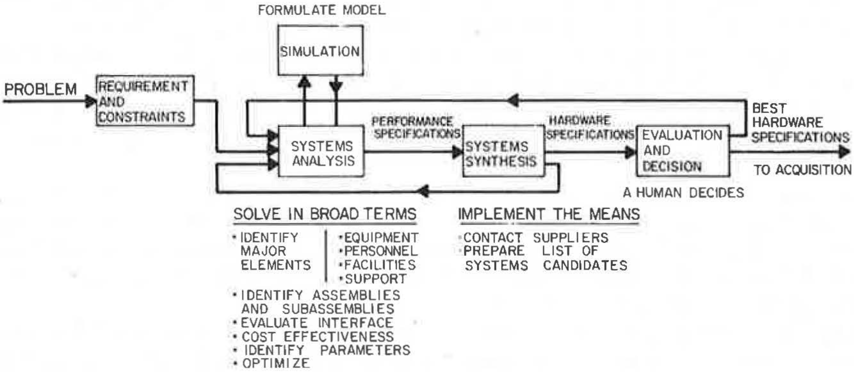


Figure 2. System design procedure.

are not realized until an iterative procedure is established that subjects all aspects of the problems to a systematic "feedback" review. This process is designed to evaluate each component of the problem solution in the context of its relationship both to the total problem and to other components of the solution.

The systems engineering procedure can be divided into five stages: concept formulation, system definition, acquisition, deployment, and phaseout.

In concept formulation the tasks are to define the problem, set objectives, and conceive alternative solutions to the problem. During the concept formulation stage, the technical and economic constraints applicable to the problem are also defined. At this stage all topics are considered at the general level, and continual care is taken not to make specific decisions prematurely. The completion of this phase should result in an accurate definition of the problem and the establishment of requirements to meet the problem.

In the system definition stage, the analysis, simulation, synthesis, evaluation, and selection are performed. At this time the system components (hardware, personnel, facilities, and support) for each alternative plan are defined. These alternatives are then subject to an optimization process as shown in Figure 2. During this process the basic system requirements and constraints are subject to the iterative analyses leading to decisions. The benefit of this stage is derived from the trade-offs affecting each component for the good of the entire system. A secondary benefit of substantial importance that emerges from this stage is the quantification of system elements before making trade-offs. This process, which requires the assignment of numerical values to system elements, also demands that analysis of the system elements be performed to permit accurate quantification. As a consequence no system element is left to change.

One of the factors of particular importance to traffic control systems is the examination of the interfaces among system elements. The evaluation of each subsystem in relation to its impact and compatibility on other parts of the system is fundamental to the basic system operation. A specific example concerning the interfaces between a computer and its vehicle detectors will be covered later in this paper.

The subject of system reliability may appear in the system design procedure as both a system requirement and a system constraint. It is a universal fact that no element of the system can be considered independent of its reliability considerations. The attainment of reliability goals also strongly influences the basic configuration of the system design and will be discussed later relative to its traffic signal system implications.

The final products of the system definition phase are the hardware and software specifications that will be used for procurement. Accompanying the hardware and software specifications may be specifications or instructions that define the operational aspects of the system implementation, evaluation, maintenance, and personnel requirements.

During the acquisition phase the necessary research, development, and design work is performed leading to production and equipment availability. Depending on the hardware specifications, the acquisition phase could simply result in the procurement of existing hardware. Software development would also proceed during acquisition. As a component of the system development phase, a timetable would have been created that would define the phasing of necessary on-site construction to effect a smooth deployment of the equipment. The construction of required facilities would also proceed during the acquisition phase.

The installation, checkout, and operation of the system occur during the deployment phase. The logistic support, maintenance implementation, and system operation procedures established during system definition would be implemented at this time. Once the system becomes operational, the evaluation concepts would be utilized and the success of the system would be established.

Planning for system phaseout would probably be superficial for a traffic control system. It is necessary, however, to recognize that equipment has a finite useful lifetime and that operation with technically obsolete equipment can be a false economy. Consequently guidelines should be established to aid in determining the point at which system replacement should be considered.

It is important to note that systems engineering is an interdisciplinary activity drawing on the talents of specialists in all aspects of a problem. The economic aspect of procurement of a traffic control system is not to be minimized. The investment in a traffic control system can be a substantial one for the community with the consequence that the economic constraints may be severe. A well-developed analysis combining both economic and engineering talents is needed to ensure that the best system is provided for the available money. In addition, a traffic control system is one of the few public facilities that has the potential of paying for itself by reducing delay to the motorist.

Figure 3 shows the basic configuration of the traffic control system and its data flow. The basic elements shown include detectors, D_n, controllers, C_n, data transmission devices, DC_n, a media for the transmission of data, TM, and the computer.

Data from the detectors, D_n, are converted to a mode suitable for transmission by data convertor DC_{1n}. The data are then transmitted by the transmission medium, TM, to data convertor DC_{4n} where they are converted into a form usable by the data input function of the computer.

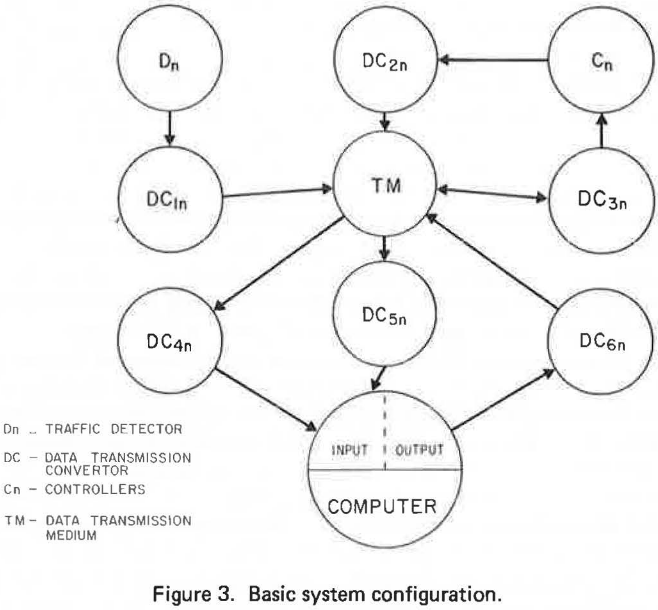


Figure 3. Basic system configuration.

The computer control decisions are converted from their form at the data output function of the computer to a form suitable for transmission to the controllers by data convertor DC_{6n}. The data then proceed via TM to data convertor DC_{3n} where they are converted to a form usable by controller C_n.

To establish that the desired control action occurs, a confirmation of the controller response is sent to convertor DC_{2n} where it is converted for transmission and proceeds back to the computer input function via TM and DC_{5n}. The only function present in this basic diagram that is not absolutely essential to system operation is the feedback of controller configuration via DC_{2n}, TM, and DC_{5n}. However, the assumption will be made that confirmation of controller synchronization is a basic system requirement.

Figures 4 and 5 show two examples of practical implementation of the basic traffic control system. The system shown in Figure 4 is characterized by a unique data transmission medium and unique data transmission convertors for each communications channel. This is probably the most direct implementation of a traffic control system and will serve as a model to apply examples of system design considerations.

Initially the detector channels will be examined. From the reliability standpoint it may be observed that the failure of any element in the detection channel will result in invalid detector data. It should also be recognized that since the detector channels are independent the failure of an element in one channel will not affect another channel. One option to enhance the reliability of the detector channel would be to provide a redundant channel; however, this raises several questions.

The basic question to be asked is, What is the effect on the total system operation if a single detector channel fails? Before this question can be answered, a determination would have to be made of the possible differences that would occur in the system operation depending on the manner in which the failure occurred. Three possible results of the channel failure can be considered:

1. The channel indicates continuous vehicle presence;
2. The channel indicates no vehicle presence; or
3. The channel shows intermittent vehicle presence with no relation to the actual traffic at the detector.

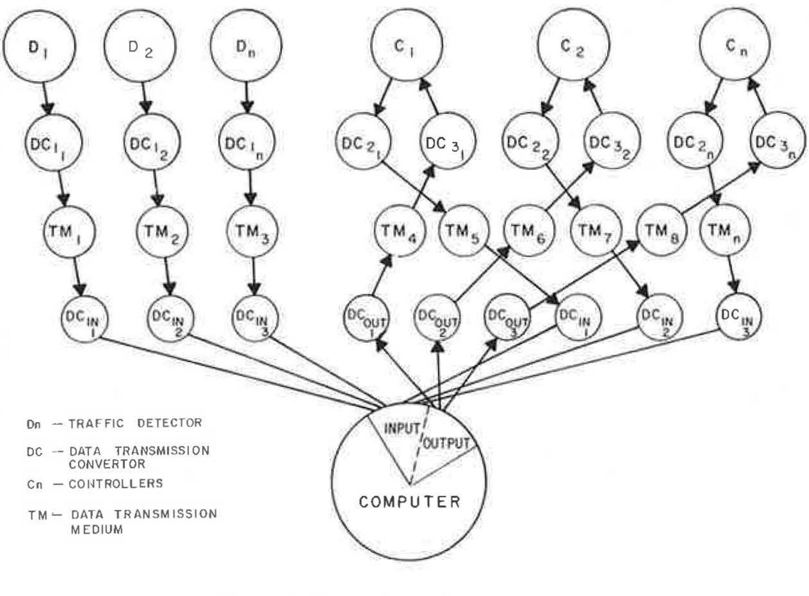


Figure 4. Example system configuration.

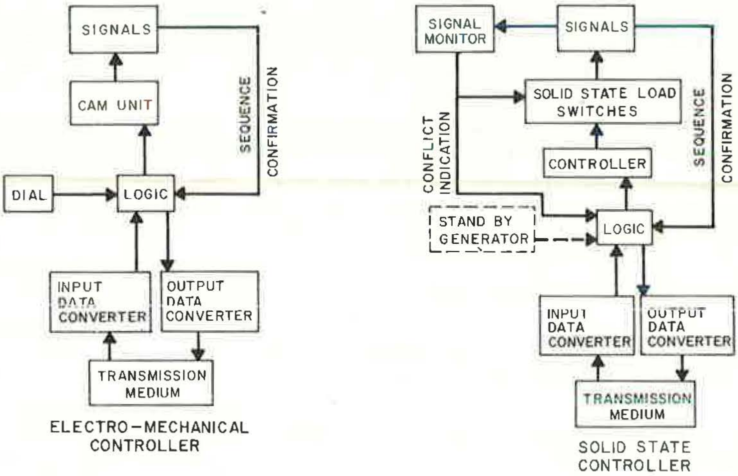


Figure 5. Pretimed controller configurations.

Because a steady failure can probably be detected and the failed detector's input data disregarded by the computer, the question of the effect on system operation of the absence of the failed detector data must be resolved. If system operation were unacceptably degraded by the absence of a single detector, then a vote might be given to the consideration of a redundant channel, but even here another question emerges: How long can degraded system operation be tolerated compared to the expected time required to repair the failed channel? If the degraded performance cannot be tolerated until the channel is repaired, another vote might be given for redundancy. Still other questions remain: How frequently (what is the mean time between failures) can a detector channel be expected to fail? What is the probability that the failure will be of a continuous type that can be more easily detected? Of course the very basic question of whether automatic detection of a failed detector can be initially provided requires resolution.

Consideration of the detector channel that failed intermittently would require additional analysis and will heavily involve consideration of a means of detecting this phenomenon.

If the evidence weighed in favor of considering redundant detector channels, an entire new analysis process would have to begin that would evaluate whether the entire channel should be redundant, whether just the least reliable element of the channel should be redundant, or whether the channel could be made satisfactorily reliable by redesign or selection of more reliable components.

The intent of the foregoing description is to demonstrate the iterative thought process intrinsic with system design procedures; it was not intended to analyze exhaustively a hypothetical problem.

A second system problem using the same detector channel can be examined based on actual data. This is the problem of detector data accuracy. Because detector data provide the real-time input descriptive of traffic flow, it is reasonable to assume that errors in this input will result in control errors unless the means are provided to compensate for the errors. For example, assume that a loop-detector is the detection element and the parameter to be examined for accuracy is the duration of time a vehicle is within the detector's zone of detection. To develop this example the following assumptions will be made:

1. The loop is 6 ft square and is located in the center of a 12-ft wide traffic lane;
2. The vehicle is 18 ft long;
3. The vehicle is traveling at a speed of 30 mph (44 fps); and
4. The duration of time that the vehicle is over the loop is given by

$$PW = \frac{L + 1}{v}$$

where

PW = pulse width in seconds,
L = vehicle length in feet,
l = dimension of loop parallel to vehicle motion in feet, and
v = vehicle speed in feet per second.

Therefore

$$PW = \frac{18 + 6}{44} = 0.545 \text{ sec}$$

Now let us examine the sources of error to this calculated value:

1. Assume a random error of ±2 percent in basic detection accuracy;	±2 percent
2. Assume an error of ±12 percent due to the change in relationship of the vehicle to the loop's electromagnetic field as the relationship of the vehicle's centerline to the loop's centerline changes;	±12 percent
3. Assume an error of ±5 percent due to the recovery time of the loop detector (if two vehicles are following closely the detector generally does not fully recover quickly enough, and the pulse width accorded the following vehicle is decreased);	±5 percent
4. Assume an error due to the drift of the detector with respect to operating temperature of ± 25 percent. (Note: This is a wide variable among loop detectors and varies from a few percent to nearly 100 percent.)	±25 percent
Total accumulate error =	±44 percent

Thus, if tolerances are additive, the actual pulse width produced by the loop detectors may lie between 0.305 and 0.785 sec (i.e., 0.545 sec ± 0.240 sec) and should be carefully evaluated with respect to system requirements. Further, the specifications for this or any other system element should reflect the accuracy requirements of the system.

Two additional sources of error exist in the detector channel—the transmission medium and the data convertors. The source of error due to the transmission medium can generally be disregarded for traffic control purposes.

Two common types of data conversion elements are relays and frequency division multiplex equipment. Relays appear to be the most elementary of devices, but even a simple relay can be a source of error. The relay error in accurately duplicating a pulse width occurs for two reasons. First the drop-out time of a relay is generally greater than the pull-in time. The drop-out time will also be increased by a factor of two to three if the relay coil is shunted by a diode for transient reduction purposes. Second, the response time of a relay changes with respect to its coil resistance, which is a function of temperature. Because the characteristics of relays vary widely with the size and type of relay, the errors due to relay characteristics may be insignificant: however, the errors should be known and evaluated with respect to the system requirements.

Frequency division multiplex equipment can possess both excellent and poor response characteristics depending on its design. In general, frequency division multiplex equipment that has a very narrow bandwidth (100 Hz or less) tends to have very sluggish response times that may disqualify it as an element in a pulse-width-dependent transmission channel.

A third example of system analysis can be shown by comparing two methods of implementing controller elements in a traffic signal system. Figure 5 shows an electromechanical controller and a completely solid-state controller operating under system control.

In both examples the fixed intervals in the signal sequence will be timed by the controller. Pulses generated by the computer will be used to terminate the green intervals, thus providing the means of regulating the cycle, split, and offset of the controller.

A feedback path is also provided in both controllers to verify that the controller is in synchronism with the computer commands.

The logic element interfaces the data convertors with the controller and also switches the controller (or cam unit) to the dial or standby generator in the event of loss of synchronism or upon receipt of a command to begin standby operation. During standby operation the signal timing will be established entirely by the dial unit or the controller and standby generator.

The most noticeable difference between the two configurations is the addition of a signal monitor element and a standby generator element in the solid-state example. The signal monitor is required in response to an assumed system requirement that no component failure should result in dangerous, conflicting signal indications. The mechanical construction of the cam unit makes this requirement intrinsic with the electromechanical controller. However, because the solid-state load-switching devices possess the possibility of failing in either a conducting or nonconducting state, a means must be provided to detect a failure in this element that would result in dangerous signal indications.

If equal reliability of the two controllers is to be achieved, then the combined reliability of the solid-state load switches must equal the reliability of the cam unit. Although solid-state devices are generally more reliable than electromechanical devices, the system engineer cannot be satisfied with broad generalities but must examine all elements and their relationships to the extent that the available data permit.

A similar analysis can be made of the dial unit, the standby generator, and the solid-state controller. In the electromechanical controller the dial unit, by its design, can remain in synchronism with inputs from the computer. The dial unit in conjunction with its cam unit also provides the same function of the solid-state controller, solid-state load switches, standby generator, and signal monitor. Consequently, because there are more elements in the solid-state configuration, their individual probabilities of failure must be higher so that their combined reliability equals the electromechanical system.

It should be noted that the standby generator might be incorporated into the solid-state controller. If this element ceases to be uniquely identifiable, the system reliability consideration becomes one of determining the effect, if any, of the reliability of the controller due to the added standby components.

The intent of this discussion is not to imply that electromechanical devices are more (or less) reliable than solid-state devices. The intent is to emphasize that only by rigorous analysis of all components and elements for a proposed signal system can an optimum design be attained.

REFERENCES

1. Reliability Stress and Failure Rate Data for Electronic Equipment. MIL-HDBK-217A, 1965.
2. Shooman, M. L. Probabilistic Reliability. McGraw-Hill, 1968.
3. Stern, S. Traffic Flow Data Acquisition Using Magnetic-Loop Vehicle Detectors. Highway Research Record 154, 1966, pp. 38-52.

LEGIBILITY AND BRIGHTNESS IN SIGN DESIGN

Bernard Adler, AIL, Division of Cutler-Hammer, Inc.; and
Arthur L. Straub, Clarkson College of Technology

An important but neglected aspect of sign design is the choice of letter heights to satisfy nighttime legibility requirements. In choosing letter heights, the fundamental relationship of brightness and legibility must be taken into account. Sign brightness is a function of many factors including sign material and position, road alignment, and vehicle and headlight characteristics. A computer program was developed that incorporates these factors and determines sign brightness as a function of road distance. The distance at which the sign must be first legible is used in conjunction with the computed brightness and published empirical data relating brightness to legibility to calculate required letter heights. Minimum letter height requirements for road distances up to 2,000 ft are presented. The cases reported include a straight road, high and low headlight beams, six sign positions, four horizontal alignments, and four vertical alignments. For nighttime legibility, it was found that required letter heights are much larger than the 50-ft-per-in. rule indicates. Because of the widely varying sign brightness found in actual roadway conditions, each sign should be treated individually as a separate design problem.

•IT is evident that, for the near future at least, the conventional highway sign will remain the principal means of transmitting information to the highway user. Increasing demands to satisfy traffic operating problems make it essential to optimize all aspects of sign design. This paper is concerned with an important but neglected aspect of sign design—the choice of letter heights to satisfy night legibility requirements.

In order for a highway sign to fulfill its purpose, its message must be legible under both daytime and nighttime conditions. At night, under typical rural conditions, with no fixed sign lighting, a sign is illuminated only by the car's headlights. Just as for any other object falling within the headlight beam, the luminance or brightness of a highway sign is a function of its position and reflectivity, the road alignment, and the position of the car on the road. In a rural area, sign brightness varies greatly. In an urban situation, where electric power is more readily available, the sign may be internally or externally illuminated and the brightness can be maintained at higher and more uniform levels. However, whether the sign is illuminated by fixed sources or by headlights, the resulting brightness, as seen by the driver, determines the sign's legibility.

Allen et al. (1) studied the relationship between sign luminance and legibility distance (the distance at which a sign can be read for a given letter height, as a function of brightness of the letter) and empirically determined a functional relationship between the two. This important relationship is shown in Figure 1. The curve is an overall average of results for medium ambient illumination without headlight glare and for low ambient illumination with and without headlight glare, for both dark legends on light backgrounds and light legends on dark backgrounds. It should be noted that, in order to obtain legibility equal to or better than 50 ft of legibility per inch of letter height (the commonly accepted design value for daylight operations), a luminance value of more than 5 ft-lamberts is required. If the brightness falls much below 5 ft-lamberts, the night legibility drops

Sponsored by Committee on Traffic Control Devices and presented at the 50th Annual Meeting.

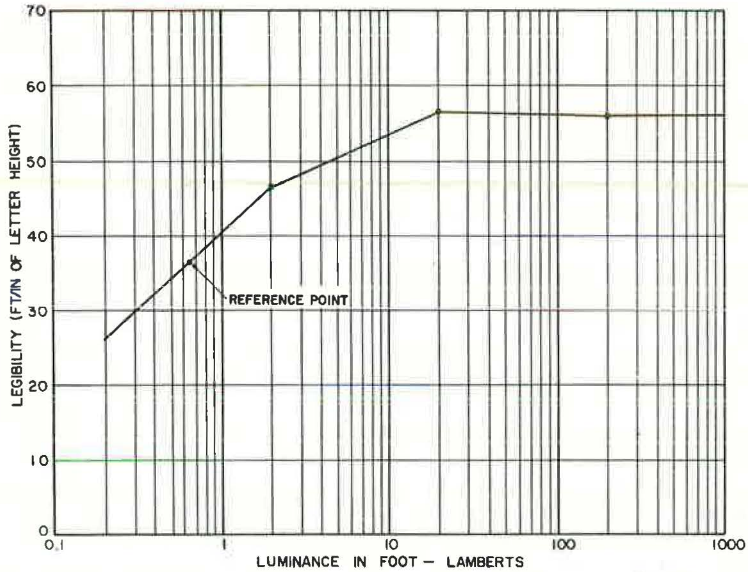


Figure 1. Legibility as a function of luminance.

far below the 50-ft-per-in. value. For many situations the preferred range is from 10 to 20 ft-lamberts. Much higher sign brightnesses are required in areas subject to high ambient illumination (as in an urban area), or where glare sources are present. A complete discussion of these factors is given by Allen et al. (1).

Many signs on our highways have a night brightness much less than 5 ft-lamberts at the point at which their messages are intended to be read. For those signs having low brightness, the commonly used 50-ft-per-in. rule is not valid, and hence many signs may not be legible at the distance assumed by the designer. The Manual for Signing and Pavement Marking of the National System of Interstate and Defense Highways (2) and the Manual on Uniform Traffic Control Devices for Streets and Highways (3) do not account for this brightness-legibility relationship.

Widespread use of retroreflective sign material has resulted in signs that are much brighter than those produced by nonreflectorized surfaces and other diffuse objects in the driver's field of view. These bright signs can result in nighttime performance that, in some cases, approaches that of good daytime conditions. It is very significant to recognize, however, that, as seen by the driver under night roadway conditions, reflective materials in common use today provide a luminance range of from less than 0.1 ft-lambert to more than 100 ft-lamberts. Wide ranges of brightness are due not only to differences in reflective properties of the material itself but primarily to wide ranges in illumination from the headlights and to the geometric relationships between the sign position and the roadway alignment. The relationship of these factors to the brightness of signs can be analytically determined for a wide range of conditions that are likely to occur on an actual roadway.

This paper describes the results of efforts to tie together two fundamental relationships concerning reflectorized signs: the legibility of the signs as a function of brightness and the brightness of the signs as seen by approaching drivers as a function of applicable parameters (sign material, road geometry, vehicle). The results are expressed in terms of minimum required letter heights. The approach to design assumes that the designer will treat legibility at a particular point or road section as a basic factor to be designed for and that letter height selection is one of the primary design decisions to be made. Hence, the basis for the development of a letter height design procedure is established.

The work described herein is a part of that accomplished under NCHRP Project 3-12. The final project report (4) contains a comprehensive account of the relationship of this work to the total information requirements and transmission techniques for highway users.

FACTORS AFFECTING SIGN BRIGHTNESS

The major factors involved in determining nighttime brightness at the driver's eye are the sign, the road, and the vehicle.

The sign factor has two subdivisions: (a) material, which establishes photometric properties, and (b) position, which is the location of the sign with respect to the road. The sign may be in the median, overhead in the median lane, overhead in the curb lane, or on the roadside mounted at several possible lateral offsets from the edge of the highway.

The road factor deals with alignment and includes straight roads, horizontal curves with different degrees of curvature and changes in curvature, and vertical curves with different grade changes and grade lengths.

The last factor is the vehicle, which includes the headlight type, high or low beam, and the classification of the vehicle (model of car, truck, etc.) that fixes the locations of the headlights and the driver's eyes. All these factors are given in Table 1.

DEVELOPMENT OF COMPUTER PROGRAM

A general analytical method for determining the brightness of reflectorized signs for a variety of sign materials, sign positions, distances, highway alignments, and traffic conditions was first described by Straub and Allen (5). A computational program was written using Fortran IV for the IBM 360/30 computer using similar techniques to determine the brightness of reflectorized signs. The program broadens the scope of the referenced work by including many additional parameters. This program was used to derive the various relationships shown and discussed in this paper.

Sufficient computer runs were made (more than 300 in all), using representative values of the applicable parameters, to demonstrate the applicability of the method and to determine, if possible, the general trend of these relationships; Figure 2 is one example of the results. A field investigation of actual brightness was made, and the results were correlated with the predicted values. A more detailed account of the computer program and its use are given in the project final report (4) and also in a paper by King (6) included in this Record.

TABLE 1

FACTORS AFFECTING SIGN BRIGHTNESS

Sign	Road

Sign face material (photometric properties)
Position
  Lateral offset
  Vertical offset
Distance from sign to vehicle
	Vertical curves
	  Beginning grade (g_1)
	  Final grade (g_2)
	  Total grade change ($g_1 - g_2$)
	  Length of curve (L)

Road	**Vehicle**

Horizontal alignment
  Tangent
  Horizontal curves
    Intersection (deflection) angle (Δ)
    Degree of curve (D)
    Length of curve (L)
    Transition spirals
Vertical alignment
  Constant grade
    Level
    Not level
	Headlights
	  Number
	  Type
	  Arrangement
	  Location
	  Beam use (high or low)
	Driver's eye position

DETERMINATION OF REQUIRED LETTER HEIGHT

Given the computed sign brightness versus road distance information for a wide variety of sign, roadway, and vehicle conditions, the next step is to make use of the brightness-legibility relationship to determine the required minimum letter heights.

Figure 3 is one example of the results. It shows the relationship of minimum letter height as a function of the required reading distances from the sign for a straight road and a sign legend made from standard sheeting-type material commonly used on Interstate signs. In applying results to design, it is assumed that only good letter designs are used, such as standard upper and lower case modified Series E (7). It is further assumed that letters are displayed at adequate contrast ratios. The curves in Figure 3 are shown for overhead and roadside signs illuminated by high and low beams.

The basic process for developing this curve is as follows:

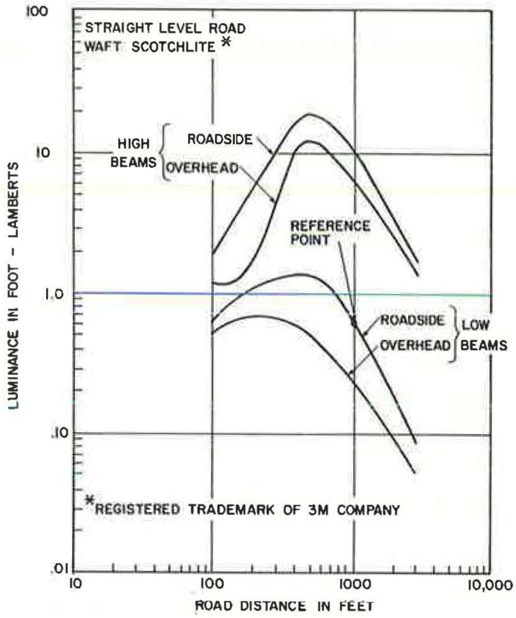


Figure 2. Sign brightness.

1. For a given road distance, find the luminance for a given sign position and beam (from data such as shown in Fig. 2). Example: for a roadside sign, low beams, and a 1,000-ft road distance, read a luminance value of 0.62 ft-lambert ("reference point" on Fig. 2).
2. Using the luminance found in step 1, use Figure 1 to find the corresponding leg-

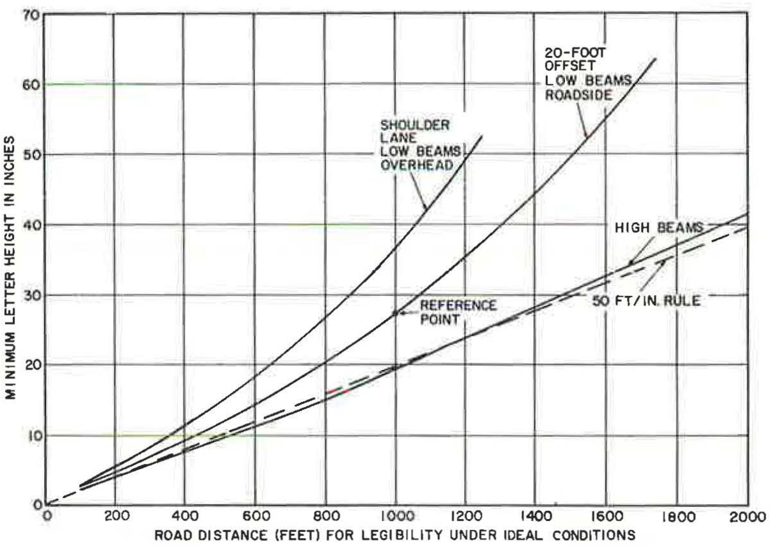


Figure 3. Minimum required letter height as a function of required legibility distances for a straight, level road.

ibility factor. Example: for 0.62 ft-lambert, read a legibility factor of 36.5 ft/in. ("reference point" on Fig. 1).

3. Divide the road distance used in step 1 by the legibility factor found in step 2 to find the letter height. Example: 1000 ÷ 36.5 = 27.4 in. The point is plotted in Figure 3 ("reference point"). This is the minimum letter height for the sign message to be legible at 1,000 ft for a car approaching a roadside-mounted sign using low beams.

4. Steps 1, 2, and 3 are repeated as required for other road distances so that a curve can be plotted to show a general relationship for a roadside sign illuminated by low beams. The same basic process, using appropriate data, was used to determine all other curves shown in this paper relating minimum letter height to road distance.

In Figure 3, the curve shown for "roadside" is for legibility at the center of a 10- by 20-ft ground-mounted sign with its left edge 10 ft from the pavement edge and its bottom 7 ft above the pavement. The curve shown for "overhead" is for legibility at the center of a 10-ft high overhead sign mounted with its bottom 17 ft above the pavement over the right-hand lane. For reference and comparative purposes, the commonly used rule of thumb, 50 ft of legibility per inch of letter height, is also plotted in Figure 3. Figure 4 shows the sign positions together with others studied in this project.

The road distance must be specified to apply this technique to a particular problem. By using techniques reported elsewhere (8, 9), an analysis of roadway and expected traffic parameters can be made to determine the distance required for the driver to process the information received from a given highway sign and to perform the required driving maneuvers safely and comfortably before reaching the decision point. This distance determines the position of the last possible point at which the information must be transmitted to an approaching driver. When transformed into the roadway length and added to the previously determined distance, message reading time (a function of sign message length and complexity) determines the position of the first point at which the sign must be legible to the driver. Between these two points is the zone within which the message must be received. From the standpoint of legibility design, the roadway distance from the sign to the first point (the point farther from the sign) is the more critical.

The following example illustrates this new approach to letter height design. Assume that an analysis of traffic maneuvering requirements for a tangent section has indicated that a sign needs to be first legible at a point 800 ft upstream from a proposed sign

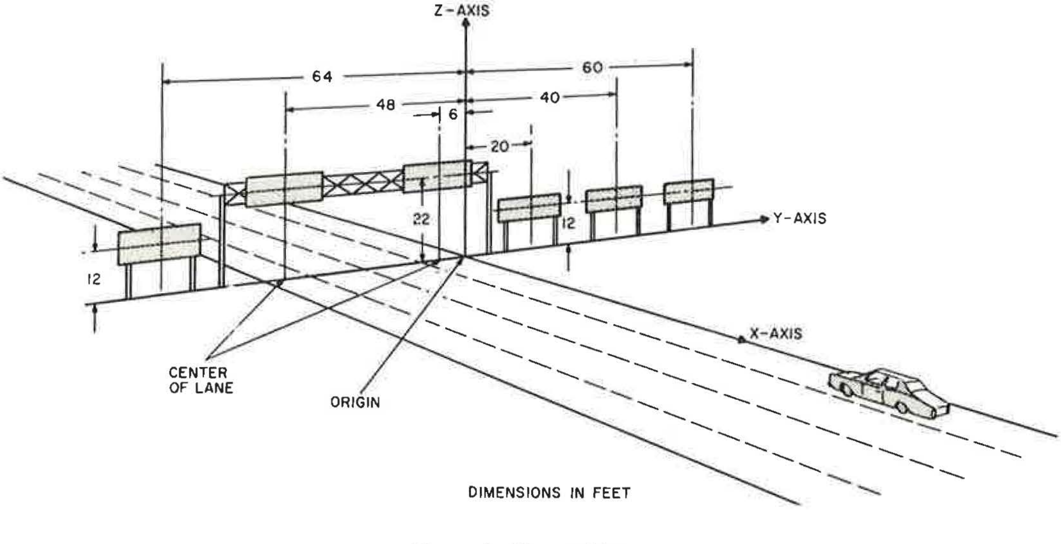


Figure 4. Sign positions.

location. Also assume that low beam use predominates and that the basic design choice being made is between an overhead and a roadside sign position. Referring to Figure 3, it can be seen that, for equal legibility, the minimum letter size for an overhead sign is 27 in. and for the roadside sign is 20 in. In practice, if a nonstandard size happened to be indicated, the designer would consider the next larger standard letter size (7). The choice of which is the better sign position would depend on economic considerations and on other design considerations to be discussed later. It is emphasized here, however, that, from the standpoint of equal legibility, the different sign positions require different letter sizes to allow for the different brightness.

For the preceding example, if a 16-in. letter height were used (based on the 50-ft-per-in. rule), the first point of legibility would be at 540 ft for the overhead sign and 650 ft for the roadside sign instead of the required 800 ft. If this fact were not recognized by the sign designer, this reduced legibility (because of reduced brightness) could lead to serious operating problems.

HEAD-LAMP BEAM USE

As can be seen from Figure 3, the curve for high beams closely approximates the 50-ft-per-in. curve shown for reference. Under high-beam illumination, both the overhead and roadside sign positions require letter heights that are nearly equal to each other; hence, only one curve is drawn. Under high-beam illumination, the legibility of the signs closely approximates acceptable daytime performance.

Although vehicles are equipped with both high- and low-beam headlight systems, however, indications are that most vehicles are operated at night using low beams. This is true even for relatively low-volume, rural, Interstate divided-highway alignments. A study in South Dakota (10) reported that 67 percent of all motorists traveling the Interstate study section were using their low beams when first sighted. A later study (11), conducted throughout the United States on both two- and four-lane roads, indicated that for a sample of over 23,000 vehicles observed under open-road conditions less than 25 percent were using high beams.

Therefore, for the purpose of designing reflectorized signs, low-beam operations must be assumed to predominate. One reservation to this statement should be kept in mind. Hare and Hemion (11) stated that "There are marked variations in beam usage habits of drivers from area to area in the United States." Thus, the designer must keep local conditions in mind before deciding on a "design beam."

The additive effects of other vehicles in the traffic stream (as they might increase the brightness of a sign as it would appear to a given driver) was the subject of a special study (4). The total additive effects are surprisingly small (because of the larger divergence angles from the other vehicles' head lamps) and, of course, cannot be counted on to occur during off-peak hours. The net result is that the design condition should be considered as a single vehicle operating on low beams.

EFFECT OF SIGN POSITION

The analysis was made at the center of a sign 20 ft wide and 10 ft high, which was faced with material considered as commonly used reflective sheeting. Six sign positions were used in this study (Fig. 4). The 20-ft offset sign is the standard ground-mounted sign. The 40- and 60-ft offset signs represent signs displaced from the highway by 30 and 50 ft respectively. The curb lane overhead sign is the standard, and the median lane overhead sign is mounted over the fourth lane of an eight-lane divided highway, with the bottom of these signs 17 ft above the pavement. The median sign is placed with its right edge 6 ft to the left of the median lane and the bottom of the sign 7 ft above the pavement. The approaching car is in the right-hand lane and the head lamps are on low beam.

Figure 5 shows the minimum required letter height curve for each of the sign positions on a straight, level road. It is noted that the letter height requirements for the 20-, 40-, and 60-ft offset signs are nearly the same, but distinctly greater than the 50-ft-per-in. rule. The median and overhead signs require very large (and impractical)

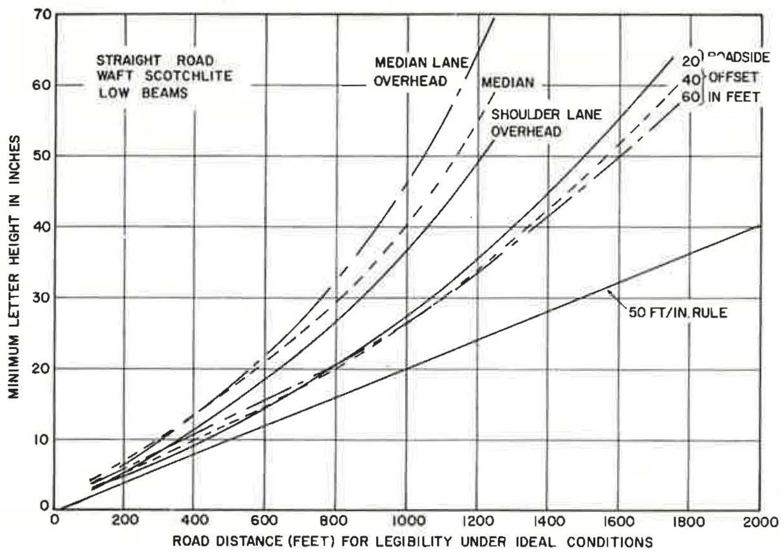


Figure 5. Effect of sign position on letter height—straight, level road.

letter sizes, especially at greater road distances, if reflectorization alone is to provide the necessary brightness.

EFFECT OF ALIGNMENT

Figure 6 shows some of the effects of horizontal curvature on the minimum required letter height for a sign offset 30 ft from the edge of the highway pavement (the center is

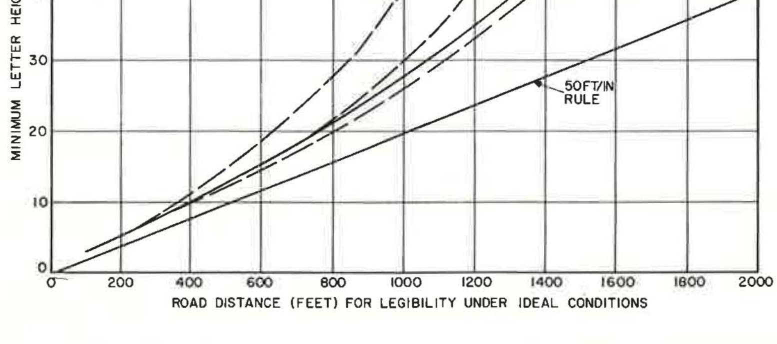


Figure 6. Effect of changes in approach horizontal alignment on letter height for a 30-ft offset sign.

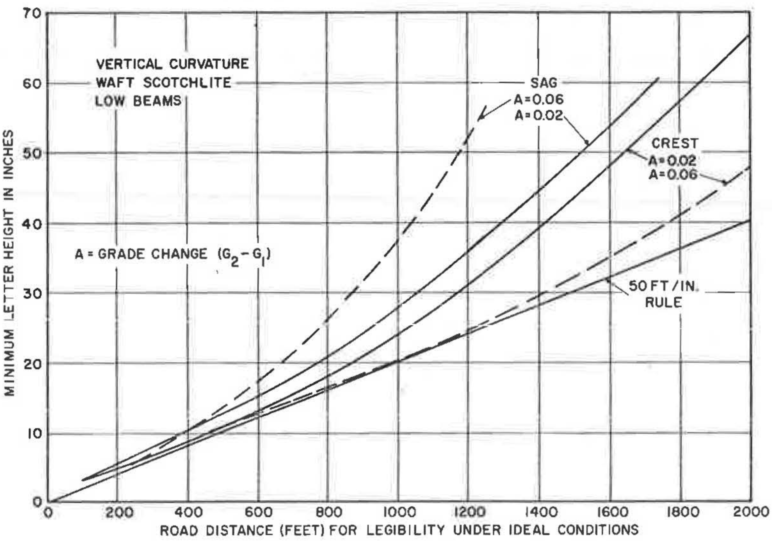


Figure 7. Effect of changes in approach vertical alignment on letter height for a 30-ft offset sign.

40 ft from the edge of the pavement). The plots are for a road curving to the right and show the effect of degree of curvature (D) and deflection angle (Δ) as a car using low beams approaches. Although not shown, the graphs for left curvature are similar in shape but show slightly greater letter height requirements.

In all cases larger letter sizes are required than those given by the 50-ft-per-in. rule. The effect is especially pronounced for the longer, sharper curve (D = 4 and Δ = 40); for example, a 40-in. letter height is required for legibility at 1,000-ft road distance, instead of 20 in. as given by the rule.

Figure 7 shows some of the effects of vertical alignment on minimum letter heights. Again the approaching car is using low beams. For these curves, as well as for the horizontal curves, the sign is offset 30 ft from the pavement edge and is located at the end of the road curvature. Figure 7 shows the results of two values of total grade change for both crest and sag curves. In each case, the recommended minimum length of curve for a design speed of 70 mph was used in the calculations (12). The effect of vertical curvature on letter size can be seen from the graph. As the curvature becomes greater, grade change increases and the letter-height requirements for the sag curve are increased. At the same time the letter heights required for a crest curve decrease. The sign at the end of the crest curve with a grade change of 0.06 requires minimum letter heights very nearly following the 50-ft-per-in. rule.

DESIGN CONSIDERATIONS

In this paper the relationship between sign brightness and sign legibility has been emphasized. Other major factors, such as the choice of legend and the limits on sign location to satisfy operating conditions, are beyond the scope of this paper. It is obvious that total sign design must take into account many factors in addition to legibility at night. However, attention is focused again on the choices a designer would have in dealing with legibility design.

Several examples have been cited in which larger letter sizes are called for to satisfy night legibility requirements. One choice available to the designer is simply to use the larger sizes needed. Larger letters would require larger sign panels, which in turn yields higher costs. For many situations, the very large sizes are completely impractical to use and other choices become mandatory.

The designer must seek another way to increase sign brightness and hence to decrease the needed size. At problem locations a more efficient (i.e., brighter) reflectorized material might be selected. If a trial sign location is likely to result in low brightness, the designer could seek another location that would serve traffic needs just as well and also provide an adequately bright sign. For example, he could avoid sign locations at the end of sag vertical curves, when possible, and use crests more often.

When reflectorization alone cannot provide the brightness and legibility required, the designer can provide the needed solution by using fixed artificial illumination, either internal or external. The availability of power and maintenance costs may preclude this as a final choice, but if brightness levels can be maintained at sufficient levels artificially (say at 10 to 20 ft-lamberts), the resulting legibility will approach daytime conditions regardless of location problems associated with reflectorized signs. For the example used previously, if sufficient artificial illumination would be provided for the overhead sign, the 16-in. letter height would provide the 800-ft legibility distance needed.

If a single sign location provides questionable night legibility, the designer can consider repeating the sign at more than one location.

These and other choices are available to the designer in considering solutions to providing adequate night legibility. The basic process would be to begin with roadway geometry and traffic operating requirements. The designer would select a trial sign location, determine trial size requirements, check on restraints and adequacies, seek alternative solutions, evaluate economics, etc., in an iterative process. Only then can a solution be found that is acceptable in providing the legibility needed for the operating conditions being designed for.

In very congested areas it may be found that satisfactory solutions using signs alone (whether under daytime or nighttime conditions) cannot be found. In such situations signs can be used extensivly, but additional technology will be required to provide supplementary driver aid systems. A complete discussion of driver aid systems is found elsewhere (4).

An important point to stress is that, for the reasonably near future, signs will play an increasingly important role in traffic operations. Because of wide variations in the legibility of signs that are used under nighttime conditions, each sign should be treated as an individual design problem. To be responsive to the actual conditions, the designer must take into account the specifics of alignment, positions, etc., appropriate for each sign.

ACUITY AND OTHER LIMITING FACTORS

Of considerable significance is whether the legibility data described by Allen et al. (1), which are the bases of results presented herein, can be applied for drivers with impaired vision. Visual acuity is a function of the angle subtended by the smallest discernible detail. The median driver has a visual acuity of 20/20, which is also the average of the observers used in Allen's study. Therefore, using Allen's results to satisfy legibility requirements implies satisfaction for at least 50 percent of the drivers on the road. If a greater percentage is to be included, drivers with lower visual acuities must be considered. The fifth percentile driver has a visual acuity of 20/70 (13). Because empirical results (like those of Allen) are lacking for drivers with impaired vision, the effect of reduced acuity on legibility distances can only be estimated from a consideration of the geometry of the visual angles. Because small angle tangents vary linearly with angles, a straight-line relationship between acuity and letter height is assumed. On this basis, the 20/70 driver requires letter heights that are 3.5 times those of the median driver. Therefore, for the example used previously, the overhead sign would require letter heights of 3.5 × 27 in. or 94.5 in., and the roadside sign would require letter heights of 3.5 × 20 in. or 70 in. for low-beam illumination. The revised values of letter height should then be considered in the overall sign dimensions, and the computer program must be rerun to verify brightness and in turn letter heights for the new sign in an iterative process until letter height, sign dimension, and brightness agreement is reached. These letter height values, even though extremely large, would still not satisfy

100 percent of the driving population. The matter of visual acuity, of course, also affects vision under daytime conditions. This represents an extremely serious problem for a small segment of the driving population.

In addition to the factors covered in this paper, several others also affect the brightness of reflectorized signs. Some of these are badly aimed headlights, changes in voltage in the lighting circuits, aging of sign materials, and transmissivity (loss of light caused by atmospheric attenuation). These factors were studied under NCHRP Project 3-12 (4), but the results are not included in this paper because of space limitations. In most cases, reduced brightness results in the need for greater letter heights than those indicated by the ideal conditions shown on graphs in this paper.

One final factor should be mentioned in considering the adequacy of signs for nighttime conditions—target value or sign visibility. The driver must have his attention drawn to the sign that he is to read before he can read it; i.e., he must select this particular signal source over all the other signal sources competing for his attention at the particular moment. The lead time required between the last point at which the sign should be detected and the point of beginning legibility cannot be determined unequivocally. It depends on the complexity of the task to which the driver is attending and on the number of competing sources. A qualitative evaluation must be made for every individual location and the proposed sign design checked for adequacy of target value. A paper by Forbes et al. (14) gives a suggested procedure for predicting sign visibility that can be used for this evaluation.

When required nighttime brightness can be defined for target value, the analytical method of determining brightness of reflectorized signs previously described can be used to predict conditions at a specific proposed sign location.

CONCLUSION

An analysis of the approach to sign design detailed in this paper clearly indicates that serious deficiencies in nighttime legibility can occur if uniform letter sizes are arbitrarily adhered to or if simplified rules of thumb (such as 50 ft of legibility per inch of letter height) are used universally without regard to specific site conditions and brightness. This is particularly true for reflectorized signs.

Relationships developed in this paper establish a new approach to the design for night legibility. To be responsive to the needs of nighttime legibility, the designer must account for the relationship of sign brightness to legibility, especially for signs of low brightness. The graphs of minimum letter heights presented here show the general requirements that typify modern Interstate road alignments. In general, to account for night legibility, signs must be made larger and/or brighter.

The graphs of minimum letter heights are based on "ideal" conditions (new, clean signs, clear atmosphere, normal vision, and so forth) to account for conditions actually found on the road. Further allowance must be made for such factors as visual acuities less than 20/20 and for diminished sign brightness because of material aging, dirt, dew, and atmospheric attenuation.

As stated in the introduction, the relationships of brightness to legibility used in the development of this paper are based on overall average results for medium and low ambient illumination. Refinements should be developed to account for requirements in areas of high nighttime ambient illumination (for example, urban areas). In general, however, higher sign brightnesses are required in areas of higher ambient illumination and in areas subject to glare.

Because of widely varying brightness conditions, each sign should be treated as a separate design problem.

REFERENCES

1. Allen, T. M., Dyer, F. N., Smith, G. M., and Janson, M. H. Luminance Requirements for Illuminated Signs. Highway Research Record 179, 1967, pp. 16-37.
2. Manual for Signing and Pavement Marking of the National System of Interstate and Defense Highways. American Association of State Highway Officials, Washington, 1970.

3. Manual on Uniform Traffic Control Devices for Streets and Highways. U. S. Dept. of Commerce, Bureau of Public Roads, 1961.
4. King, G. F., and Lunenfeld, H. Development of Information Requirements and Transmission Techniques for Highway Users. Final Report, NCHRP Project 3-12/1, AIL, A Division of Cutler-Hammer, Dec. 1969.
5. Straub, A. L., and Allen, T. M. Sign Brightness in Relation to Position, Distance, and Reflectorization. HRB Bull. 146, 1956, pp. 13-44.
6. King, G. F. Determination of Sign Letter Size Requirements for Night Legibility by Computer Simulation. Presented at 50th Annual Meeting and included in this Record.
7. Standard Alphabets for Highway Signs. U. S. Dept. of Commerce, Bureau of Public Roads, June 1966.
8. Mitchell, A., and Forbes, T. W. Design of Sign Letter Sizes. Proc. ASCE, Vol. 68, No. 1, 1942, p. 104.
9. Research on Road Traffic. Road Research Laboratory, London, 1965.
10. Cheeseman, M. R., and Voss, W. T. Motor Vehicle Headlight Beam Usage on a Section of Interstate Highway 90. Highway Research Record 164, 1967, pp. 17-20.
11. Hare, C. T., and Hemion, R. H. Headlight Beam Usage on U. S. Highways. Southwest Research Institute, San Antonio, Dec. 1968.
12. A Policy on Geometric Design of Rural Highways. American Association of State Highway Officials, Washington, 1965.
13. Schmidt, I., and Connolly, P. L. Visual Considerations of Man, the Vehicle, and the Highway. Society of Automotive Engineers, Publ. SP-279, March 1966.
14. Forbes, T. W., Pain, R. F., Joyce, R. P., and Fry, J. P. Jr. Color and Brightness Factors in Simulated and Full-Scale Traffic Sign Visibility. Highway Research Record 216, 1968, pp. 55-65.

DETERMINATION OF SIGN LETTER SIZE REQUIREMENTS FOR NIGHT LEGIBILITY BY COMPUTER SIMULATION

Gerhart F. King, AIL, Division of Cutler-Hammer, Inc.

A previously developed computer simulation program was used to examine the adequacy of letter sizes on existing signing along a 20-mile stretch of Interstate highway. A total of 63 signs were analyzed in detail for minimum required legibility under three sets of conditions: daylight, high-beam illumination, and low-beam illumination. Although the standards of signing, according to accepted traffic engineering criteria and according to adherence to pertinent signing manuals, were found to be above average, computer analysis revealed considerable deficiencies. These deficiencies were most pronounced for low-beam illumination and were especially severe for overhead signs. The analysis revealed letter size deficiencies for 27 percent of all signs for daylight, 35 percent for high-beam illumination, and 49 percent for low-beam illumination. For guide signs, the corresponding figures were 39, 44, and 63 percent. Only 1 of 13 overhead signs was found to be adequate for low-beam illumination. Considerable variation in required letter height was found even between identical signs. Analysis of these variations and of the other deficiencies revealed a disproportionate influence of apparently minor changes in approach horizontal and vertical alignment.

•PREVIOUS research (1) has shown that the prime determinant of sign legibility, expressed as feet of legibility per inch of letter height, is luminance expressed in footlamberts. An empirically derived functional relationship between these two variables is shown in Figure 1.

A computer program has been developed that will determine the brightness of reflectorized signs under various conditions of illumination. [The use of an earlier version of this program to investigate parametric relationships concerning sign brightness is reported in this Record by Adler and Straub (4).] The program, documented in detail elsewhere (2) and described briefly in the Appendix, follows the general procedure described by Straub and Allen (3). As presently constituted the program permits the insertion into computer storage of an actual highway alignment, taken from construction plans, and the determination of the brightness of any sign at any point along this alignment for any specified type of vehicle approaching in any specified lane. This paper is a report of the application of this computer analysis to the signing currently in place on a stretch of Interstate highway.

TEST SITE

The tests were carried out on a stretch of Interstate 85 located in Durham and Orange Counties, North Carolina, near the city of Durham. Table 1 gives a general summary of pertinent site characteristics.

The road in question is a modern Interstate highway running in a general east-west direction. For purposes of this analysis a 20-mile section of the eastbound roadway was selected. The original signing plans were obtained and field-checked. At the time of the field check a total of 94 signs were noted. Their distribution by type is given in Table 2. This table also shows the number and types of signs selected for analysis. The

Sponsored by Committee on Traffic Control Devices and presented at the 50th Annual Meeting.

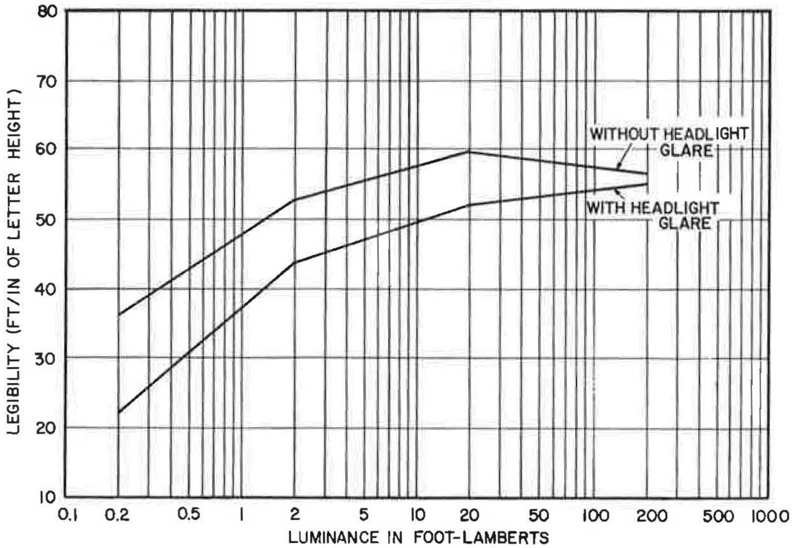


Figure 1. Legibility as a function of luminance.

signs selected for analysis were generally those contained in the original signing plans and found at the same location at the time of the field check.

A total of 63 signs were analyzed. Input data for the program were prepared from the construction plans, the signing plans, and the North Carolina signing specifications (5). The sign material characteristics used were those for WAFT Silver Scotchlite (a registered trademark of the 3M Company) after 5 years or more of exposure as supplied by the manufacturer.

Prior to the analysis the entire section was inspected, under both daytime and nighttime conditions, by an experienced traffic engineer. In his judgment the existing signing, as a whole, came up well to accepted Interstate standards. The major criticism made was that the classification of some of the interchanges as minor or intermediate, as inferred from the number and location of advance guide signs, was open to question and that some of these could equally have been signed as intermediate or major respectively.

Night inspection of the roadway also revealed that some of the demountable copy had deteriorated considerably and

TABLE 1

TEST SITE CHARACTERISTICS

Length	20.2 miles
Number of lanes	4 and 6
Lane width	12 ft
Median	30 ft and variable
AADT	18,000
Average speed	63.0 mph
Accidents/mile/year	3.52
Number of interchanges	14
Average interchange spacing	1.6 miles
Percentage of commercial traffic	28.6
Percentage of out-of-state traffic	21.3

TABLE 2

SIGNS ON I-85

Type	Analyzed	Not Analyzed	Total
Regulatory			
Do not litter	0	3	3
Speed limit	3	4	7
Reduce speed ahead	0	1	1
Warning			
Merging traffic	14	0	14
Ice on bridge[a]	0	11	11
Low clearance[b]	0	2	2
Guide			
Advanced guide	13	1	14
Exit direction	13	0	13
Exit	13	0	13
Destination and distance	2	0	2
Supplementary exit	0	1	1
Confirmatory route marker[c]	5	0	5
Miscellaneous[d]	0	8	8

[a] Folding signs, observed in closed position.
[b] Opposite-mounted at one location.
[c] One location with three shields.
[d] Includes crossroad identification and city limit signs.

had lost a considerable portion of its reflectivity. Because the research budget did not allow actual field testing of the material, the computed results, using standard photometric properties of aged Scotchlite, are probably on the high side.

RESULTS OF COMPUTER ANALYSIS

The results of the computer analysis are given in Table 3. This table shows the sign number (arbitrarily assigned to identify computer output), the type of sign, and the type of mounting in columns 1 to 3. These data were taken from the original signing plans, and field-checked for accuracy.

The required legibility distance for each sign was computed and is given in column 4. This is the distance at which a sign must become legible in order for the reading to be completed before the sign passes outside the cone of normal vision. The procedure used in computing this distance is one developed by Mitchell and Forbes (6). Figure 2 shows the geometry used in determining this distance, which is represented by AC. It is made up of two components: BC is the actual reading distance and is a function of reading time and speed; AB is the distance from the point where the sign leaves the cone of clear vision (designated by θ and usually taken as 10 deg) to the sign. The case shown is for a three-lane roadway with the driver's eye position considered to be two-thirds over in the third lane. S is the lateral offset of the sign from the edge of the road, and W is the width of the sign. A speed of 60 mph was used in the calculations, and the Road Research Laboratory formula—$t = 0.31N + 1.94$, where t = reading time in seconds and N = number of familiar words on the sign—was used to compute reading time (7).

Columns 5 through 10 show the results of the computer simulation. Apparent brightness of the sign to a driver located 400, 600, and 800 ft away, for both high- and low-beam illumination, was taken from the computer run. The figure shown is the value computed for the center of the sign.

Columns 11 through 15 show the existing letter height for various types of copy. When more than one value is shown for letter size, different lines in the sign had different letter heights specified.

Finally, columns 16 through 18 show the computed required minimum letter height. Three separate values are shown. The first is daytime letter height obtained by applying the 50-ft-per-in. rule to the required legibility distance. The other two represent minimum letter heights for nighttime use for both high- and low-beam illumination. These values were obtained by converting luminance into unit reading distances using the data shown in Figure 1. For high beams the curve "with headlight glare" was used,

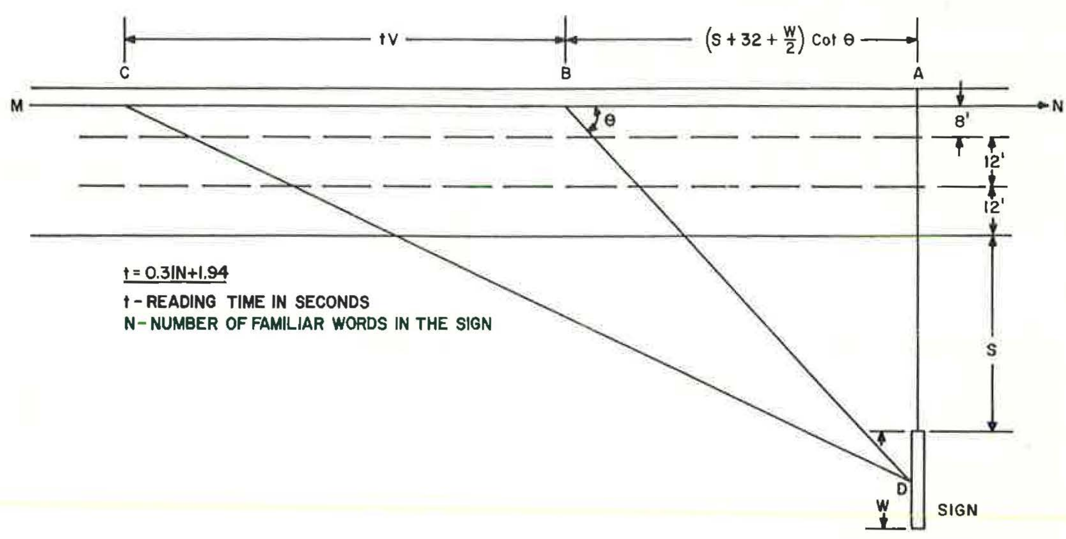


Figure 2. Determination of legibility distance.

TABLE 3

EXISTING AND REQUIRED LETTER SIZE

Sign No. (1)	Type[a] (2)	Type of Mounting[b] (3)	Required Legibility Distance (ft) (4)	Apparent Brightness (ft-lambert) 400 ft High (5)	400 ft Low (6)	600 ft High (7)	600 ft Low (8)	800 ft High (9)	800 ft Low (10)	Existing Letter Height (in.) U.C. (11)	L.C. (12)	Caps (13)	Numerals (14)	Numerals in Shield (15)	Required Letter Height (in.) Day (50 ft/in.) (16)	High Beam (17)	Low Beam (18)
1	ED	GR	460	6.90	0.96	9.77	0.89	8.19	0.71	16	12				10	12	10
2	EX	GR	410	3.22	0.48	1.93	0.27	1.07	0.17			12			10	10	10
3	MT	GR	460	11.53	2.04	12.56	3.65	8.21	4.90			8			10	10	8
4	EX	OH	490	4.69	0.24	5.28	0.18	4.25	0.14					12	10	12	13.33
5	TT	OH	540	6.50	0.26	6.15	0.20	4.95	0.17	16	12			12	12	12	16
6	CR	GR	430	8.52	0.95	10.46	0.86	8.52	0.63			7		12	10	10	10
7	SL	GR	430	3.20	0.47	3.65	0.31	4.44	0.31			8	16		10	10	12
8	DD	GR	600	3.79	0.64	3.03	0.53	2.15	0.40	13.33	10		13.33		12	13.33	16
9	AG	GR	630	4.95	0.70	7.64	0.75	7.34	0.61	16	12	10	15		13.33	13.33	16
10	ED	GR	570	4.17	0.68	4.08	0.63	3.55	0.54	16	12				12	13.33	13.33
11	EX	GR	410	7.44	0.86	9.46	0.82	6.66	0.62			12			10	10	10
12	MT	GR	460	8.29	0.93	10.01	0.85	6.92	0.64			8			10	10	10
13	CR	GR	430	15.69	0.72	7.65	0.25	2.05	0.10			7		12	10	10	10
14	AG	GR	690	15.34	0.87	13.45	0.47	3.21	0.37	16	12	10	15	12	16	16	18
15	ED	GR	630	4.32	0.56	6.90	0.57	6.42	0.49	16	12			12	13.33	13.33	16
16	EX	GR	410	2.65	0.37	1.33	0.21	0.69	0.12			12			10	10	12
17	MT	GR	460	3.35	0.56	1.98	0.38	1.02	0.17			8			10	12	12
18	MT	GR	460	3.36	0.56	1.98	0.38	1.32	0.25			8			10	12	12
19	CR	GR	430	8.20	0.89	9.72	0.81	6.95	0.50			7		12	10	10	10
20	SL	GR	430	10.62	1.41	13.32	1.91	9.64	2.91			8	16		10	10	10
21	DD	GR	610	5.44	0.71	8.17	0.73	7.34	0.58	13.33	10		13.33		13.33	13.33	13.33
22	AG	GR	690	10.46	0.84	8.40	0.57	4.16	0.22	16	12	10	16, 15		16	16	18
23	ED	GR	630	5.64	0.72	8.63	0.77	8.77	0.97	16	12	16	16		13.33	13.33	16
24	EX	GR	410	7.57	0.87	9.84	0.83	8.20	0.62			12			10	10	10
25	MT	GR	460	8.46	0.95	10.40	0.86	8.48	0.64			8			10	10	10
26	CR	GR	430	8.03	0.86	9.81	0.80	8.18	0.60			7		12	10	10	10
27	SL	GR	430	7.68	0.85	9.50	0.80	6.19	0.43			8	16		10	10	10
28	AG	GR	800	4.46	0.64	6.52	0.63	6.04	0.49	16	12	16, 12, 10	16, 15		16	18	24
29	MT	GR	460	5.52	0.68	3.33	0.46	2.07	0.35			8			10	10	12
30	EX	OH	590	5.95	0.26	5.74	0.32	3.21	0.30	16	12	16, 12	16		12	13.33	16
31	TT	OH	510	2.37	0.18	5.22	0.18	4.98	0.30	16	12	12		12	12	12	—
32	AG	GR	590	6.49	0.96	9.47	0.90	8.04	0.72	16	12	10	15		12	12	13.33
33	MT	GR	440	16.40	0.87	9.88	0.31	4.53	0.28			8			10	10	12
34	MT	GR	440	15.53	0.64	6.66	0.25	1.98	0.10			8			10	10	10
35	ED	GR	560	11.67	0.68	5.49	0.18	1.49	0.08	16	12	10			12	12	16
36	ED	GR	530	5.35	0.73	8.07	0.77	7.51	0.61	16	12				12	12	12
37	EX	GR	400	6.11	1.17	8.31	2.36	7.90	2.73			12			8	10	10
38	MT	GR	440	1.67	0.30	1.47	0.25	1.17	0.27			8			10	12	12
39	RL	GR	480	6.27	0.80	8.82	0.80	7.86	0.62	16	12	10			10	10	12
40	ED	GR	450	6.16	0.78	8.68	0.78	7.78	0.61	16	12				10	10	10
41	EX	GR	400	7.31	0.84	9.44	0.81	8.15	0.62			12			8	10	10
42	MT	GR	440	8.16	0.91	10.21	0.84	8.42	0.64			8			10	10	10
43	TT	GR	740	6.20	0.82	10.73	1.24	9.17	1.75	16	12	15, 10	15	15	16	16	16
44	ED	GR	520	21.98	1.37	16.88	0.54	11.43	0.46	16	12	10			12	12	12
45	EX	OH	490	5.26	0.25	3.64	0.15	2.38	0.10	16	12				10	12	13.33
46	ED	OH	650	4.71	0.39	3.44	0.19	2.35	0.11	16	12	15, 10	15	15	13.33	16	—
47	EX	OH	680	2.85	0.38	2.99	0.25	2.41	0.16			7		15	16	16	20
48	MT	GR	440	8.41	0.96	9.25	0.75	4.79	0.31			8			10	10	10
49	EX	GR	580	3.60	0.58	4.79	0.64	5.97	0.59	16	12	15		15	12	13.33	13.33
50	MT	GR	440	4.43	0.73	3.45	0.59	2.87	0.50			8			10	10	10
51	CR	GR	590	6.15	0.69	8.36	0.68	6.10	0.55			7		12	12	12	13.33
52	SG	GR	530	9.92	2.46	9.46	0.86	3.72	0.24	16	12	10			12	12	12
53	EX	OH	510	2.90	0.40	5.01	0.35	5.09	0.33	16	12	10			12	12	13.33
54	ED	OH	650	5.05	0.39	5.37	0.23	4.75	0.22	16	12	12, 10	15	15	13.33	16	18
55	TT	OH	490	5.17	0.24	5.33	0.17	4.55	0.15			18			10	12	13.33
56	MT	GR	440	6.12	0.63	7.88	0.62	7.18	0.52			8			10	10	10
57	EX	OH	580	4.95	0.26	5.37	0.18	2.82	0.20	16	12	15		15	12	13.33	—
58	ED	OH	630	3.89	0.22	3.49	0.14	2.55	0.19	16	12	12, 10	10		13.33	16	—
59	TT	OH	500	2.12	0.18	2.30	0.13	2.29	0.17			18			10	16	—
60	EX	GR	560	8.00	1.66	16.44	0.83	3.77	0.13	16	12	12			12	12	12
61	MT	GR	440	11.86	2.16	16.33	3.93	9.13	0.31			8			10	10	10
62	AG	GR	590	5.20	0.65	5.85	0.46	3.33	0.21	16	12	15, 10	15	15	12	13.33	16
63	ED	GR	540	8.90	1.55	11.96	3.40	7.54	5.30	16	12	15, 10		15	12	12	12

[a] ED = exit direction; EX = exit; MT = merging traffic; TT = "through traffic"; CR = confirmatory route marker assembly; SL = speed limit; DD = destination and distance; AG = advance guide; RL = "right lane"; SG = supplementary guide.

[b] GR = ground-mounted; OH = overhead-mounted.

whereas for low beams the curve "without headlight glare" was used. The luminance value was obtained by straight-line interpolation for the required reading distance from the computer results. The letter heights, computed by dividing required reading distance, are shown to the next highest "standard" size. [Standard size refers to sizes normally commercially available, for both cutout and demountable letters. These are also the sizes listed in the Standard Alphabets (8).] It should be noted that if the "no headlight glare" curve of Figure 1 had been used to compute these values, some of the required letter heights for high-beam illumination would have been actually smaller than those computed for daytime use.

DISCUSSION OF RESULTS

Comparison of the computed letter size requirements with the letter sizes actually required was made in terms of standard letter sizes. Table 4 gives the statistical distribution of this comparison. Two separate comparisons were made, one for all signs analyzed and one for guide signs (exclusive of confirmatory route markers) only. A total of 41 signs fell into this latter category.

The "actual" size used in the comparison was the largest size letter on the sign. In cases where lowercase lettering was used in the largest size line, the height of the lowercase letter governed. It should be noted that considerable differences are encountered when the smaller "subsidiary" copy is examined. This includes such items as the cardinal direction for route markers, the exit message in advance guide or exit direction signs, and all other copy which, in accordance with the AASHO Manual (9), requires smaller letter sizes.

Figure 3 is a graphical representation of the data in Table 4; it shows a considerable degradation of performance when the illumination median goes from daylight through high beam to low beam. The cumulative distributions, plotted in Figures 4 and 5 for the two cases considered, show that 73 percent of all signs are adequate for daylight illumination, 65 percent for high beam, and 51 percent for low beam.

When guide signs only are considered, the situation is much more serious. The comparative figures for complete adequacy are 71, 56, and 37 percent respectively. Although several instances were noted where low-beam requirements were equal to, or even

TABLE 4

LETTER SIZE COMPARISON

Category		Day		High Beam		Low Beam	
		Percent	Cumulative Percent	Percent	Cumulative Percent	Percent	Cumulative Percent
All Signs							
Larger than required	+4	3.2	3.2				
	+3	4.8	8.0	6.3	6.3	3.2	3.2
	+2	3.2	11.2	1.6	7.9	7.9	11.1
	+1	34.9	46.1	33.3	41.2	27.0	38.1
Adequate	0	27.0	73.1	23.8	65.0	12.7	50.8
Smaller than required	-1	19.0	92.1	19.0	84.0	14.3	65.1
	-2	7.9	100.0	14.3	98.3	17.4	82.5
	-3			1.6	100.0	7.9	90.4
	-4						
	-5					1.6	92.0
	∞					7.9	100.0
Guide Signs							
Larger than required	+4	4.9	4.9				
	+3	—	4.9	2.4	2.4		
	+2	4.9	9.8	—	2.4	2.4	2.4
	+1	21.9	31.7	19.5	21.9	17.2	19.6
Adequate	0	39.1	70.8	34.2	56.1	17.2	36.8
Smaller than required	-1	17.2	88.0	21.9	78.0	14.6	51.4
	-2	12.2	100.0	19.5	97.5	21.9	73.3
	-3			2.4	100.0	12.2	85.5
	-4					—	85.5
	-5					2.4	87.9
	∞					12.2	100.0

Note: Excess or deficiency expressed in terms of "standard" letter sizes.

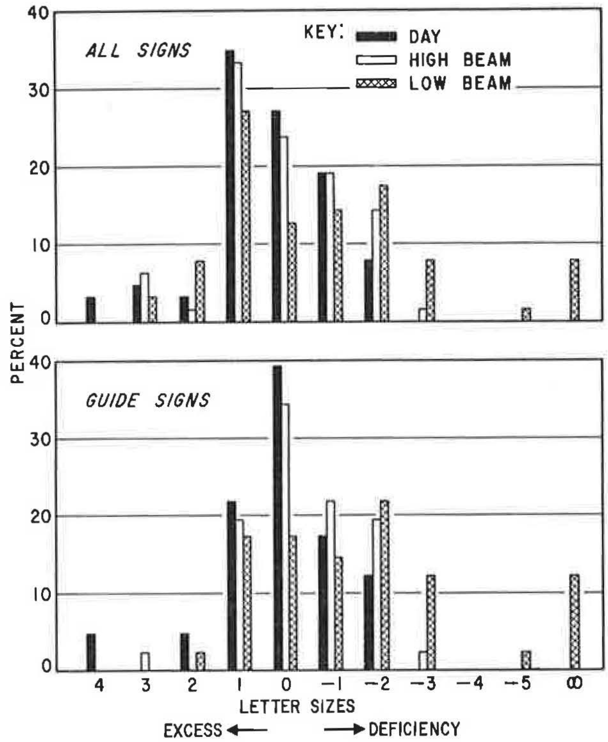


Figure 3. Letter size comparison.

smaller than, those for the other two conditions investigated, the overall impact of these figures is that low-beam illumination appears to be inadequate for signs designed under current standards. The important fact to be noted is that 5 of the 13 overhead-mounted signs were completely "illegible" under low-beam illumination; i.e., the computed brightness at the point of required first legibility was less than 0.2 ft-lambert, the lowest value for which legibility data are available. Seven of the eight overhead signs for which legibility could be computed were found to require letters two or more standard sizes higher than that required for daytime conditions.

One additional point should be made concerning the computations for overhead signs. Each sign was analyzed individually; i.e., the reading time was computed on the basis of the message on that sign alone. A strong case can be made, however, for considering an overhead sign assembly, sign bridge, or butterfly as a single message to be read in its entirety by the approaching driver. In that case the reading time would be considerably longer, the required legibility distance increased considerably, and the required letter sizes correspondingly larger. For the 13 signs included in the 5 overhead sign assemblies considered here, the average required legibility distance would increase from 560 to 780 ft or by almost 40 percent.

Table 5 gives a computation of minimum letter size under the assumption that each of the overhead assemblies is considered as a single sign. The brightness values used were the average for the signs making up the assembly; otherwise the computations are as previously described.

The increase in minimum required letter sizes is immediately apparent, amounting to at least 2 standard sizes for all conditions investigated. It can also be seen that 3 of the 5 assemblies are "illegible" under low-beam illumination, and the 2 assemblies that can be made legible can be made so only by using letter sizes that approach the limits

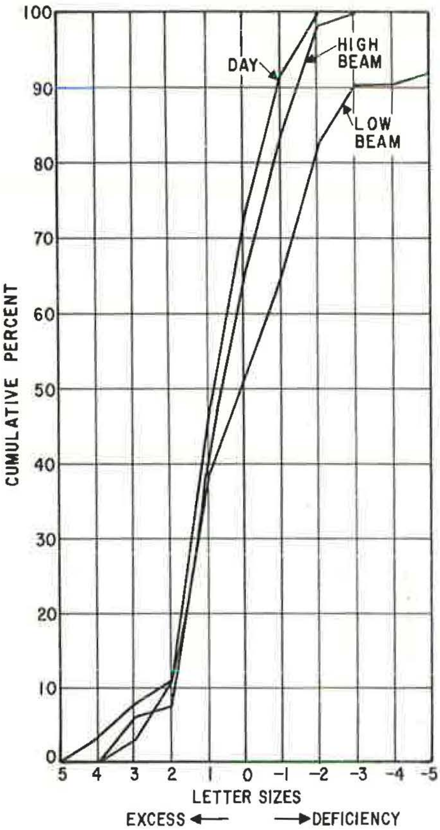


Figure 4. Letter size comparison, cumulative distribution, all signs.

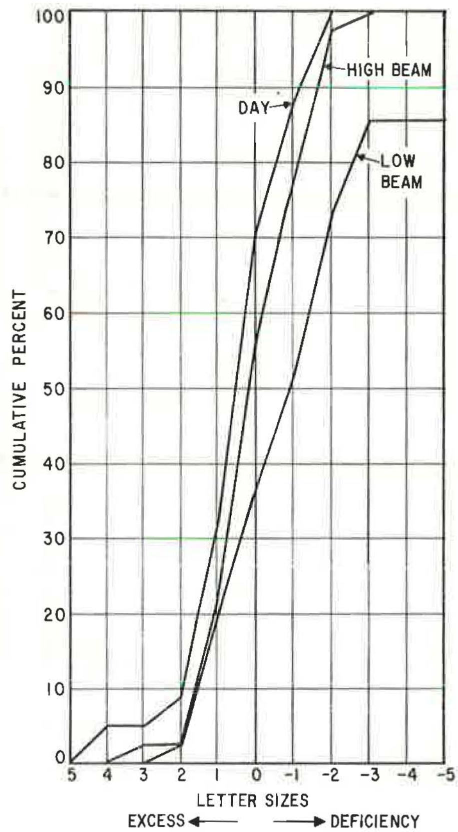


Figure 5. Letter size comparison, cumulative distribution, guide signs only.

of practical feasibility. In this connection it should be mentioned that the analysis was made while using constant sign panel sizes. The increase in panel size necessitated by the increase in letter size would result in lower luminance because of higher deflection angles as well as in increased required legibility distance as a result of moving the last reading point upstream. Both of these conditions would, in turn, result in still larger letter sizes.

TABLE 5

OVERHEAD SIGNS

Assembly	Signs	Required Legibility Distance (ft)	Existing Letter Height (in.) U.C.	L.C.	Caps	Numerals	Numerals in Shield	Required Letter Height (in.) Day (50 ft/in.)	High Beam	Low Beam
A	4, 5	630	16	12			12	13.33	13.33	—
B	30, 31	700	16	12	16, 12	16	12	16	16	20
C	45, 46, 47	930	16	12	15, 10, 7	15	15	20	24	—
D	53, 54, 55	830	16	12	18, 12, 10	15	15	18	18	22
E	57, 58, 59	800	16	12	18, 15, 12, 10	10	15	16	18	—

CASE 1. OCT. 2,1969. HIGHWAY US 70, EASTBOUND SECTION BETWEEN GREENSBORO AND DURHAM, N.C.
PROCESSED NOV. 7, 1969 13.52.10

S I G N N O. 9

STATION	OFFSET	HEIGHT	LENGTH	WIDTH	MATERIAL
15900.00	24.00	7.00	222.00	66.00	WAFT SILVER SCOT

V E H I C L E T Y P E HORZ

OFFSET	LIGHT CODE
-6.00	4

ADJ. DRIVER HEIGHT	TRANSMISSIVITY	AGE	VOLT FACTOR	MISAIM 1	MISAIM 2	MISAIM 3	MISAIM 4
3.95	1.000	5.	1.0	0.0	0.0	0.0	0.0

T O T A L L U M I N A N C E

ROAD DISTANCE	HIGH BEAM SIGN EDGE LT.-BOT.	LT.-TOP	CENTER	RT.-TOP	RT.-BOT.	LOW BEAM SIGN EDGE LT.-BOT.	LT.-TOP	CENTER	RT.-TOP	RT.-BOT.
[illegible]	0.03	0.08	0.03	0.03	0.04	0.00	0.00	0.00	[illegible]	[illegible]
[illegible]	0.02	0.02	0.02	0.03	0.03	0.00	0.00	0.00	[illegible]	[illegible]
[illegible]	[illegible]	5.43	6.02	5.49	6.30	0.38	0.27	0.39	[illegible]	[illegible]
[illegible]	[illegible]	2.19	7.34	5.52	6.67	0.80	0.91	0.61	[illegible]	[illegible]
[illegible]	11.24	7.77	7.64	4.45	5.93	0.98	0.59	0.78	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	4.95	5.00	4.59	1.23	0.60	0.78	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	2.59	1.51	3.56	0.77	3.49	0.46	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	0.78	0.59	1.00	0.69	0.21	0.44	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	0.73	0.49	0.49	0.71	0.21	0.26	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	0.64	0.55	0.56	0.47	0.23	0.21	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	0.40	0.0	0.0	0.37	0.0	0.11	[illegible]	[illegible]

10

Figure 6. Typical computer output, sign luminance.

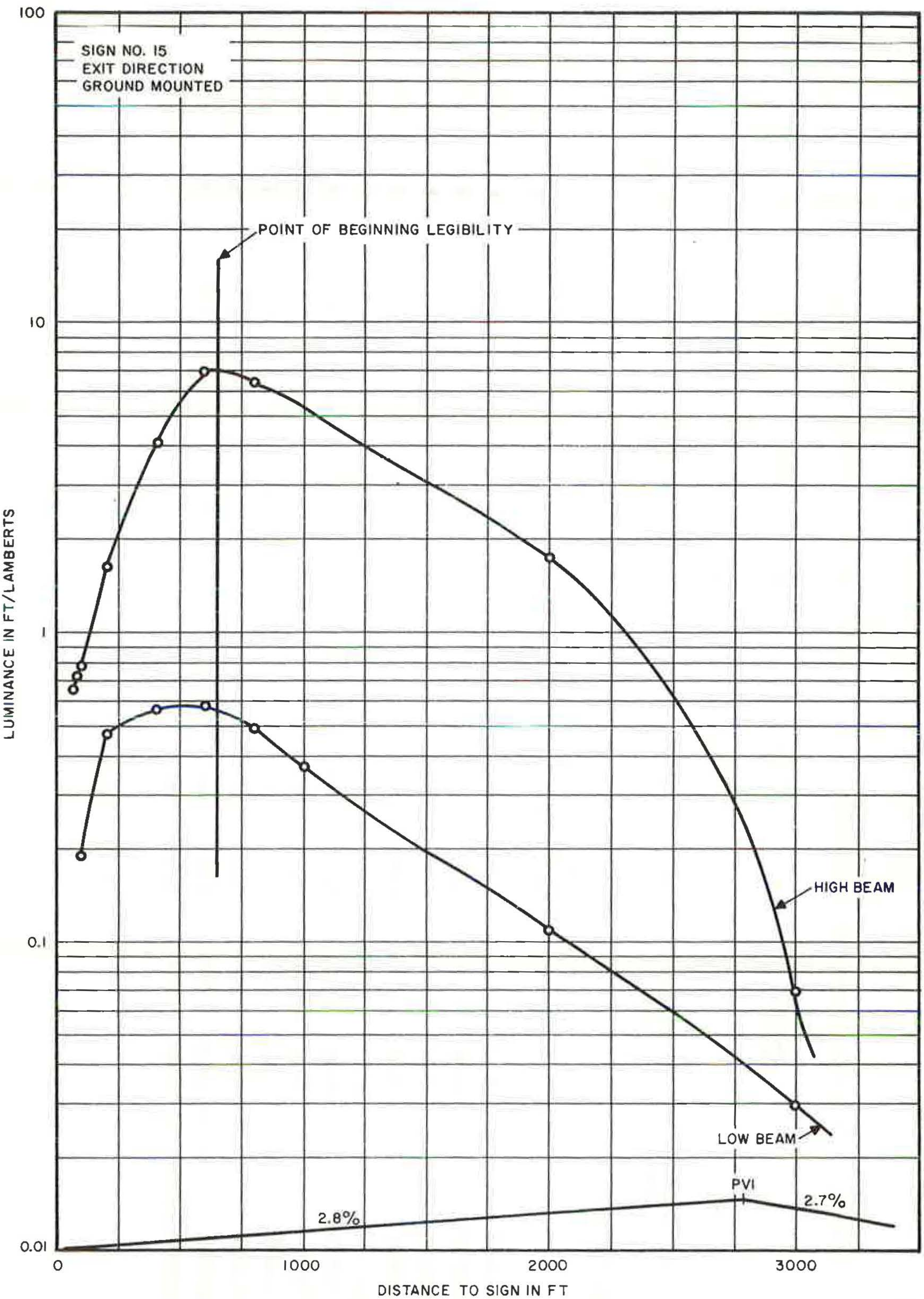


Figure 7. Luminance versus distance, sign 15.

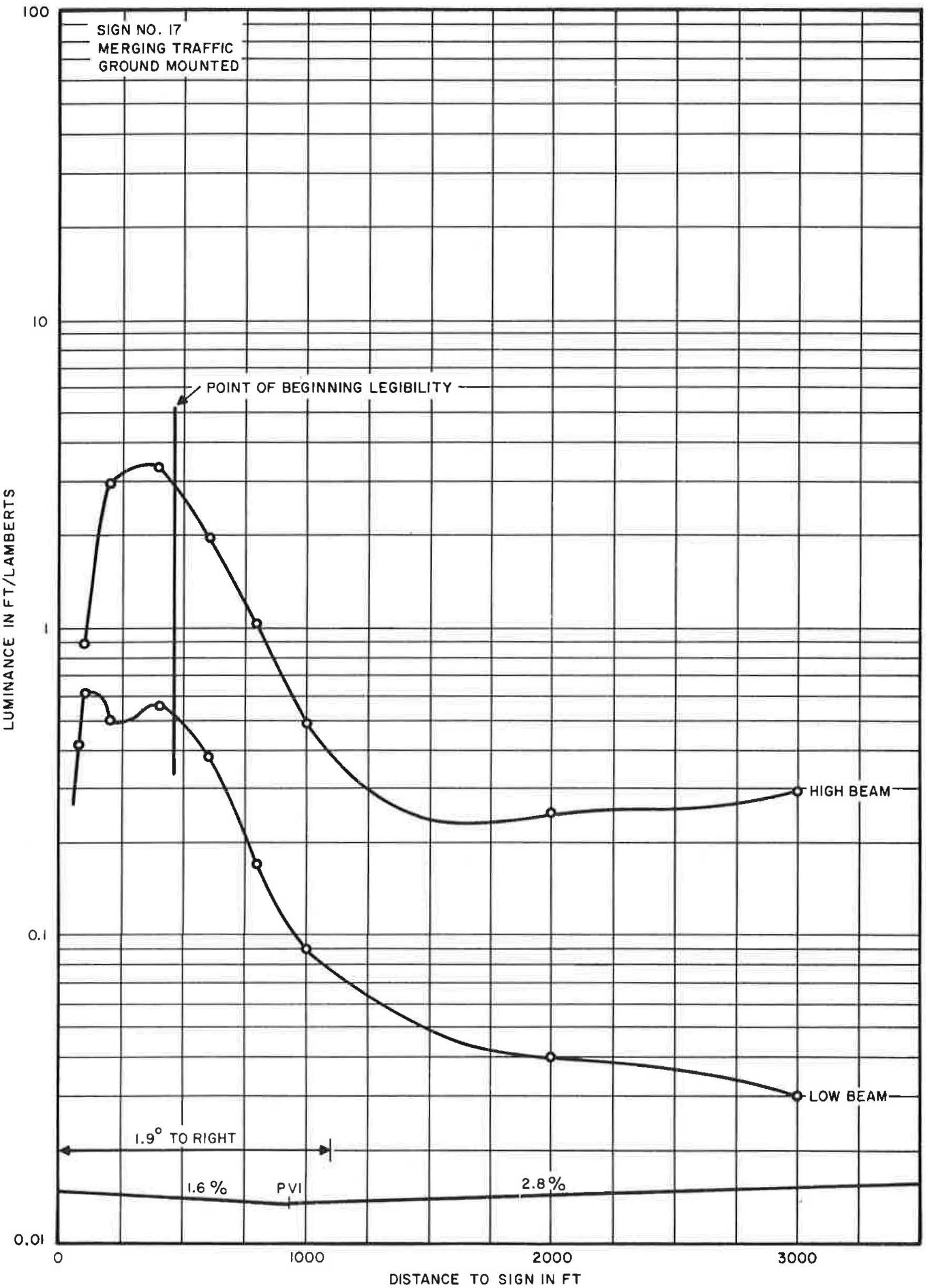


Figure 8. Luminance versus distance, sign 17.

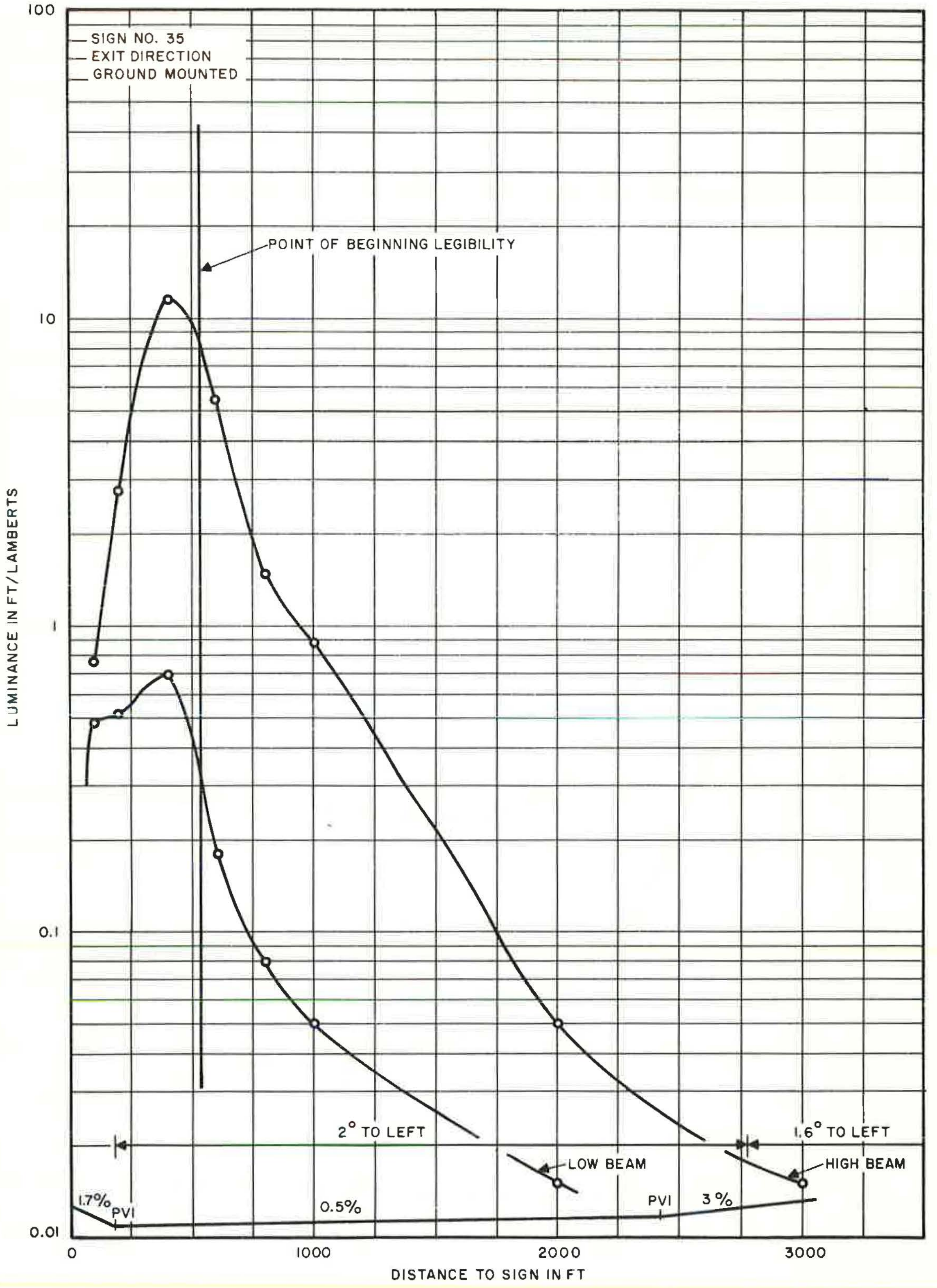


Figure 9. Luminance versus distance, sign 35.

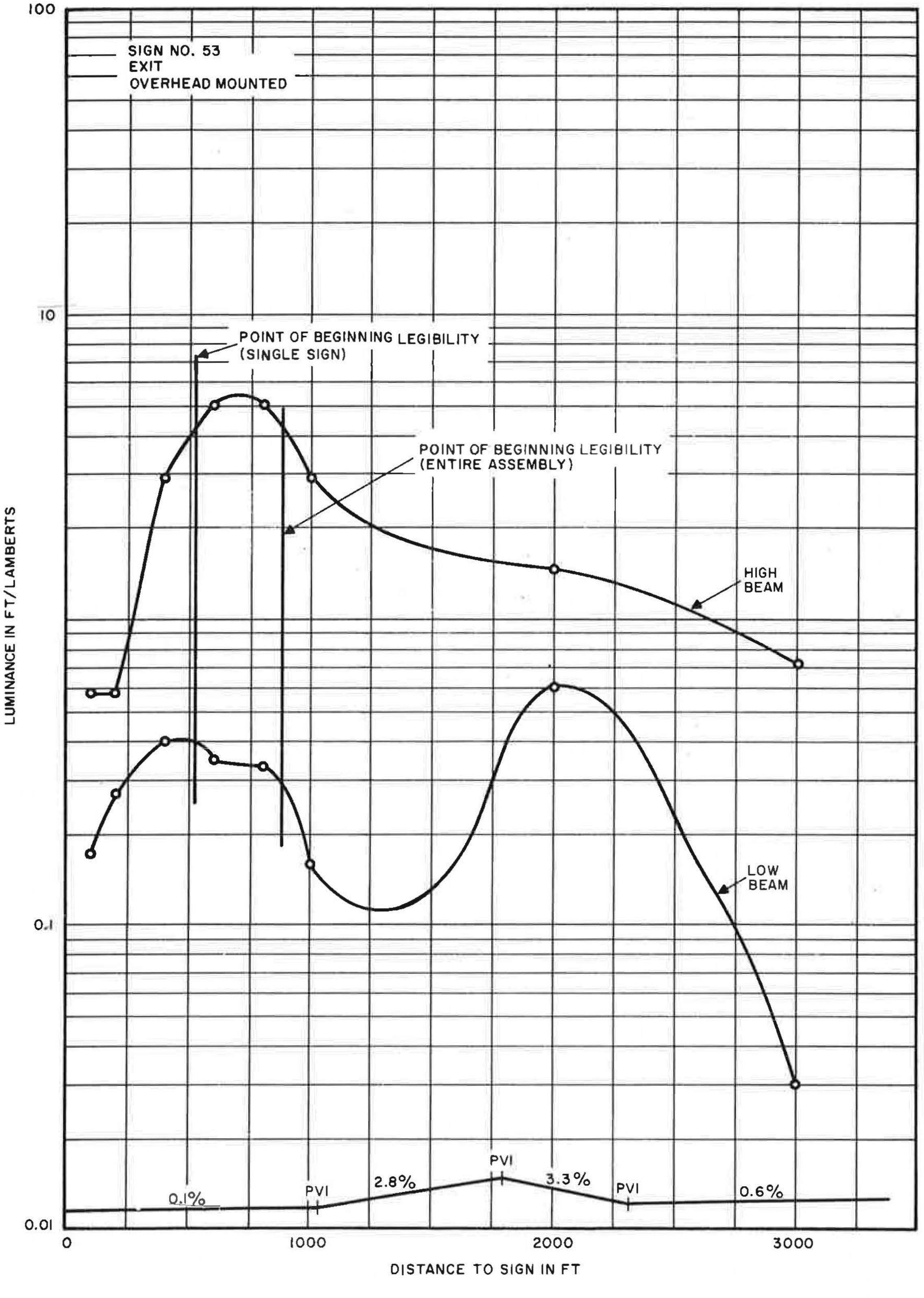


Figure 10. Luminance versus distance, sign 53.

It is also worth noting that all of the analyses were made using luminance values computed for the center of the sign. Examination of a typical computer output sheet (Fig. 6) shows that some variation exists between values for the center and values for the extremities. If worst case, instead of center, luminance had been used in the calculations, still higher computed letter heights would have been obtained.

Close examination on the data and computed results of Table 3 shows the considerable variation that exists for apparently identical signs. For instance, the 14 merging traffic signs analyzed all required 10-in. letters for daytime conditions. However, when analyzed for nighttime conditions using low-beam illumination, the analysis showed one location where 8-in. letters were required, 8 locations that required 10-in. letters, and 5 locations where 12-in. letters were required. The variation increases as sign size increases. Signs 55 and 59 are identical overhead-mounted 9- by 12-ft signs with the legend through traffic. However, analysis shows a difference of two standard sizes in the high-beam case and a difference between "legible" and "illegible" in the low-beam case.

This great variance is due to the enormous influence that apparently minor changes in horizontal and vertical alignment, especially the latter, have on the distribution of headlight illumination reaching the signs and on the resulting luminance. Figures 7 through 10 are plots of luminance versus distance from the signs for six signs used in the analysis. At the bottom of each sign the approach horizontal and vertical alignments are shown. Absence of horizontal alignment information indicates a tangent section.

Worthy of note is the smooth curve of Figure 7 that has no alignment changes but does have a continuous downgrade throughout the entire approach. Figure 8 shows a discontinuity in the low-beam curve in the area of combined vertical and horizontal alignment changes. However, these changes are not severe enough to affect the high-beam curve. Figure 9 shows an extremely sharp peak due to a horizontal curve to the left, with the sign at the beginning of the subsequent upgrade tangent section. Finally, Figure 10 shows the effect of extreme changes in vertical alignment, extreme by Interstate standards of design, on the approaches to a sign.

One final point should be made concerning these results. The signs, as analyzed, were located laterally in accordance with the standards in effect at the time of installation (1960). They are, therefore, quite close to the edge of the shoulder. A change to a 30-ft offset, as currently required, would have major effects on the computed results. Reference to Figure 2 clearly shows that any increase in S would increase distance AB and therefore distance BC. This would result in having the point of first legibility moved further upstream, requiring larger letter sizes. The effects of this increased offset were discussed in detail in a previous paper (10).

CONCLUSIONS

This case study, in applying the computer simulation program to determine requisite letter size of signs for adequate nighttime visibility, demonstrates the versatility that this approach has as a tool for the optimum design of signs. The wide variations in results noted point up the necessity of designing each sign for the exact conditions and location for which it will be used. Finally, the great variation noted between sign design for high-beam use and that for low-beam use, which, as stated, would have been even higher if the assumption of no headlight glare had been made, underlines the need to determine the prevailing headlight use. Two published studies on headlight use (11, 12) indicate that reliance should not be placed on high-beam illumination. Because the results of these computations show the inadequacy of relying on low-beam illumination for overhead-mounted signs, the conclusion seems inescapable that overhead signs require fixed illumination if they are to serve their purpose properly and effectively.

In closing this discussion, it should be emphasized that this section of highway analyzed is signed well in accord with all Manual requirements at time of installation. Furthermore, comparison of actual field conditions with the signing plans indicate that every effort has been made to correct deficiencies in signing that have become apparent since the opening of the highway. The deficiencies revealed by the present analysis are not, therefore, attributable in any way to the North Carolina highway department. Similar, or worse, results could undoubtedly have been obtained in any other jurisdiction, especially those that still use nonreflectorized sign backgrounds. The deficiencies noted represent deficiencies in the present state of the art of sign design for night legibility as reflected in the Manual and in design procedures.

REFERENCES

1. Allen, T. M., Dyer, F. M., Smith, G. M., and Janson, M. H. Luminance Requirements for Illuminated Signs. Highway Research Record 179, 1967, pp. 16-37.
2. King, G. F., and Lunenfeld, H. Development of Information Requirements and Transmission Techniques for Highway Users. Final Report on NCHRP Project 3-12, AIL, A Division of Cutler-Hammer, 1969.
3. Straub, A. L., and Allen, T. M. Sign Brightness in Relation to Position, Distance, and Reflectorization. HRB Bull. 146, 1956, pp. 13-44.
4. Adler, B., and Straub, A. L. Legibility and Brightness in Sign Design. Presented at the 50th Annual Meeting and included in this Record.
5. Sign Specification Supplement. North Carolina State Highway Commission, Jan. 1969.
6. Mitchell, A., and Forbes, T. W. Design of Sign Letter Sizes. Proc. ASCE, Vol. 68, No. 1, 1942.
7. Research on Road Traffic. Road Research Laboratory, London, 1965.
8. Standard Alphabets for Highway Signs. U. S. Dept. of Commerce, Bureau of Public Roads, 1966.
9. Manual for Signing and Pavement Marking of the National System of Interstate and Defense Highways. American Association of State Highway Officials, 1961.
10. King, G. F. Some Effects of Lateral Sign Displacement. Highway Research Record 325, 1970, pp. 15-29.
11. Hare, C. T., and Hemion, R. H. Headlight Beam Usage of U. S. Highways. Southwest Research Institute, San Antonio, Dec., 1968.
12. Cheeseman, M. R., and Voss, W. T. Motor Vehicle Headlight Beam Usage on a Section of Interstate Highway 90. Highway Research Record 164, 1967, pp. 17-20.

Appendix

PROGRAM DESCRIPTION

The computer program, called ROADSIGN, has been written to simulate a motorist driving on a road at night with headlights on and approaching a sign on the road. The program computes the luminance at the four corners and at the center of the sign as apparent to the driver of the vehicle traveling along an existing, proposed, or hypothetical highway and approaching the sign.

The simulation takes into account the following parameters that together constitute the major influences on sign legibility:

1. Distance from sign;
2. Automobile headlight output;
3. Sign material reflectance;
4. Highway alignment and profile;
5. Sign location and attitude with respect to the highway;
6. Vehicle geometry;
7. Location of driver within vehicle;
8. Location of vehicle on highway;
9. Atmospheric transmissivity;
10. Variations in vehicle voltage as they affect headlights; and
11. Variations in sign reflectance as typically caused by aging and weathering.

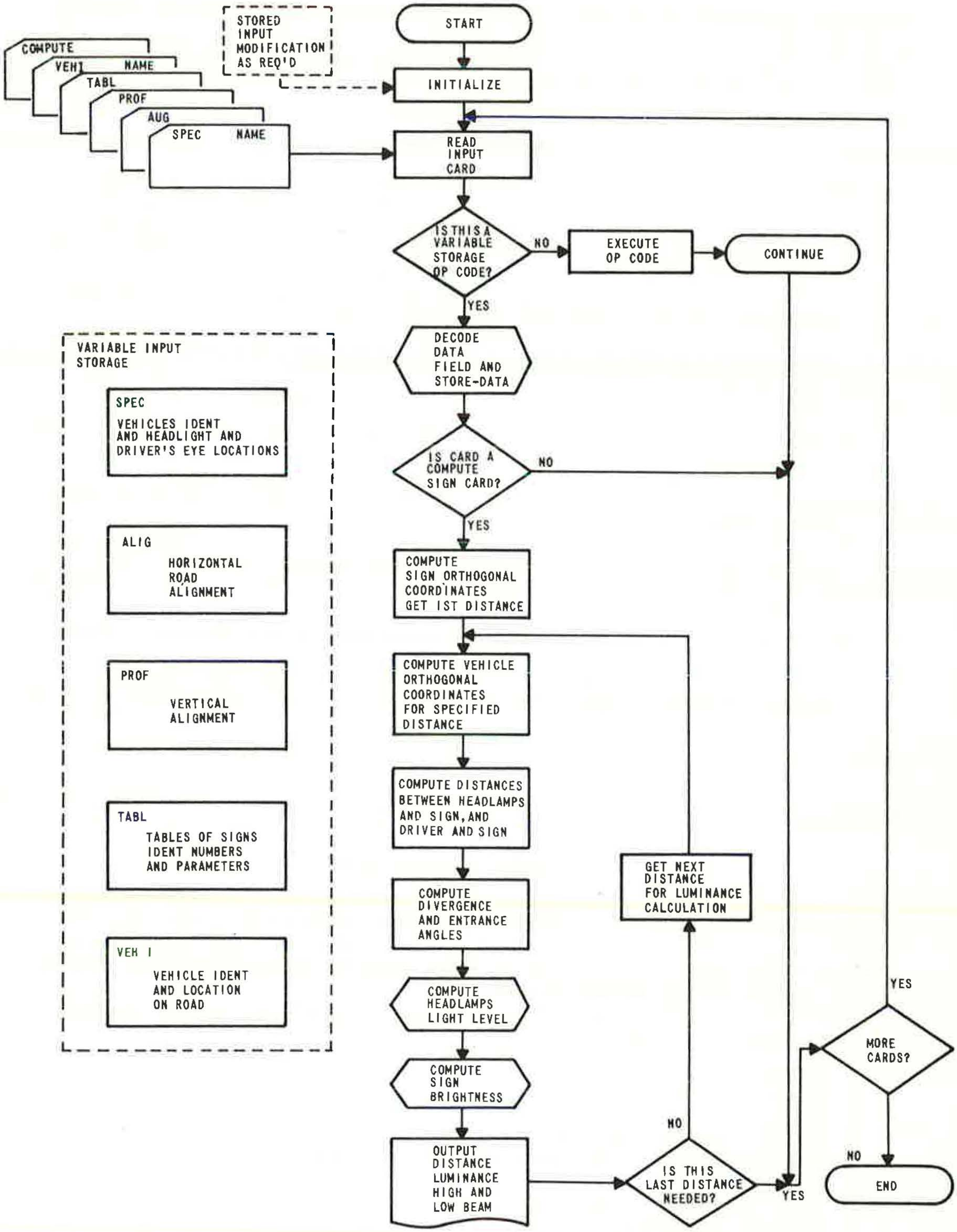


Figure 11. Program block diagram.

ROADSIGN has been designed to take into account these factors with a high degree of flexibility. In addition, in order to accommodate future expansion, the program has been structured so that additional factors not considered part of the original simulation can be incorporated without revision of the basic program structure. As currently constituted the program can handle any combination of horizontal and vertical curvature. It cannot handle spirals or other transition curves that must be approximated by circular curves. It also cannot handle crown or superelevation.

The program is written in Fortran IV to run on an IBM 360/40 or larger computer. In its present form the program will handle up to 100 signs on any length of highway containing no more than 98 horizontal and 98 vertical curves. The program is designed to handle up to 10 different vehicular configurations. The isocandle distribution tables of 5 common standard headlight types are stored internally and others can be input. The specific luminance tables for three commercially available retroreflective materials (two types of sheeting and one type of button) are also internally stored, and others may be input.

The program operates as shown in Figure 11. In general, the logic functions so as to determine the location of vehicle head lamps, driver's eyes, and sign (both center and corners) in orthogonal space. It then computes the distances between these three sets of points for each distance stored.

By using the distances computed as well as the inputs describing illumination levels, specific luminance, divergence and entrance angles, and other factors characterizing or affecting light level, the program calculates apparent sign luminance. Following each set of calculations, the program outputs results via the line printer.

Options available by means of selection of the proper control card include the choice of different types of vehicles approaching the sign of interest and/or the computation of a number of signs of the same or different configurations placed along the same highway.

The specific study reported on herein dealt with an alignment containing 29 horizontal curves and a profile containing 56 vertical curves. The sign table contained a total of 63 signs. The actual computer elapsed time required to calculate the luminance of all the 63 signs as the vehicle approached each sign from 3,000 ft to 40 ft was 12 min 47 sec. This time included reading in and checking the card inputs, storing the information, making preliminary calculations, dumping into the printer these preliminary calculations, further checking, and finally outputting 63 pages of luminance values (Fig. 6) for both high- and low-beam illumination.

EVALUATION OF REAL-TIME VISUAL INFORMATION DISPLAYS FOR URBAN FREEWAYS

Conrad L. Dudek, Texas Transportation Institute; and
Hal B. Jones, City of Austin, Texas

A questionnaire and slide presentation, designed by a multidisciplinary team, were administered for the purpose of determining driver preferences for real-time visual information displays for urban freeways. A total of 505 employees of 17 organizations in the cities of Houston and Dallas participated in the survey. Evaluation of the responses provided design inputs for the development of a real-time freeway information system. It was found that participating motorists preferred real-time information displays that were simple in nature over designs containing diagrams that orient them to the freeway and arterial streets. They also indicated a preference for unique design features, such as the use of color, to distinguish between usual and abnormal traffic conditions. In addition, the survey indicated that the motorists favored a design that explicitly distinguishes real-time visual displays from other types of freeway signing. Evaluation of symbols (circle, arrow, or bar) that could be used as part of a real-time visual display indicated no preference for any of these symbols.

• TO broaden the application of real-time freeway operations systems, the Texas Transportation Institute and the Texas Highway Department, in cooperation with the U.S. Department of Transportation, began a research project entitled Freeway Control and Information Systems. This project is an outgrowth of previous research on the Gulf Freeway in Houston that culminated in an operational freeway ramp control system (1).

One of the objectives of the project is to develop functional requirements for a freeway communications system. Toward this end, it was reasoned that the motoring public should play a major role in establishing the system design because the purpose of the system was to help fulfill their transportation needs. A questionnaire and slide presentation, designed by a multidisciplinary team, were therefore administered to obtain inputs from the motoring public.

This report discusses one major phase of the survey directed at an evaluation of driver preferences for real-time visual information displays. A large volume of results was obtained from the survey, and additional results are reported in other publications (2, 3). A total of 505 licensed drivers participated in the survey, 329 from Houston and 176 from Dallas. Some of the social and driving characteristics of the participants are summarized in Table 1.

BASIC VISUAL DISPLAYS

Four basic designs were developed to evaluate driver preferences for real-time visual information displays:

Design 1—This sign contained words and color indications to describe the traffic conditions;

Sponsored by Committee on Motorist Information Systems and presented at the 50th Annual Meeting.

TABLE 1

SOCIAL AND DRIVING CHARACTERISTICS OF PARTICIPANTS

Percentage of Participants

Sex		Age		Education	
Male	68	24 or younger	24	Grade school	4
Female	32	25 to 44	45	High school	29
		45 to 64	31	Business college	12
				Two years of college	21
				Graduated from college	34

Occupation		Received Driver Education		Years of Driving Experience	
Professional	30				
Technician	26	None	45	0 to 4	5
Clerical	22	Classroom	13	5 to 14	36
Salesworker	3	Behind the wheel	15	15 to 24	21
Craftsman	8	Classroom and behind the wheel	27	25 to 34	22
Service worker	2			35 to 44	13
Other blue collar	4			45 or older	3
Student	5				

Miles Driven per Year		Trips Via Freeway per Week	
Less than 8,000	14		
8,000 to 12,000	28	None	3
12,000 to 18,000	37	1 to 5	15
18,000 to 30,000	18	6 to 10	26
Over 30,000	3	11 to 20	39
		Over 20	17

Daily Use of Freeway for Work		Perference of Travel in Urban Areas	
Yes	70	Freeway	90
No	30	City streets	10

Design 2—This sign used only color indications to reflect the traffic conditions;

Design 3—This sign showed a diagram of the area and used illuminated color symbol indications to show traffic conditions; and

Design 4—This sign showed a diagram of the area and gave travel speeds between reference points.

There was one exception to this pattern, which will be discussed later. All of the signs were similarly designed with white letters on a green background, a red indication to describe congested conditions, and a green indication to specify normal conditions. The diagrams, when used, were illustrated in white. Travel speeds were depicted by using white numerals on a black background.

The designs were such that only one basic difference existed between Designs 1 and 2, between 2 and 3, and between 3 and 4 respectively. Consequently, the participants' choice of, for example, Design 1 over 2 would indicate a preference for the use of word messages to describe the traffic conditions. A preference for diagrams to help the motorist in orienting himself in the street system would be reflected by the selection of Designs 3 and 4. Analysis of the basic differences will help to determine the desirable characteristics of the final design.

Through the use of a slide presentation, the participants were confronted with three separate hypothetical situations: Case I, displays on the major street; Case II, displays on the frontage road; and Case III, displays on the freeway. In each of the three cases the participants were asked to evaluate sign design alternatives that would display the necessary freeway traffic information.

In the first situation, the participants were requested to assume that they were driving along a major street that ran parallel to the freeway. Their intended route was to turn right on a major street, proceed to the freeway, and then turn left onto the freeway. For some reason, the lanes of the freeway had become heavily congested. This congestion would cause extra delay in their trips if they continued to use the freeway. Changeable message signs located in advance of the major street intersections would inform the drivers of the existing condition on the freeway.

The second hypothetical situation was the same as in the previous case, except it was assumed that the respondents, as drivers, had already committed themselves to the freeway service road. By means of signs located in advance of the freeway entrance ramps, they would be informed of the traffic condition on the main lanes of the freeway.

In the third situation, the participants were asked to assume that they were driving on the freeway and were approaching the congested area. Signs on the freeway would inform them of the condition ahead.

Each of these three cases was individually presented to the participants. They were asked to rank each sign independently, giving it a rating from a low of 1 to a high of 5, according to how well it described the traffic condition to them as motorists. After each sign was individually rated for a particular case, the participants were shown a slide containing all four designs and were asked to rank these according to their preferences. Although the basic designs were similar for each of the three cases, the signs were shown in random order for each case to eliminate any bias that may have occurred from the order of presentation.

The purpose of the individual rating tests was to determine if any of the basic designs were acceptable as candidates. For example, if all designs received very low ratings, one could assume that none of the alternatives was acceptable to the participants. If some received high ratings whereas others received low ratings, one could evaluate the basic differences between the signs that were most desirable to the participants.

Ranking, on the other hand, was used as a test to determine the relative desirability of the various designs in cases of equal ratings. For example, if two designs were given equal ratings as to their abilities to communicate the appropriate messages, then the rankings would produce the relative desirability between them. Mean rankings were computed by assigning 4 points for each first choice, 3 points for each second choice, 2 points for each third choice, and 1 point for each last choice.

When ratings are given to individual items, the question arises as to what constitutes "good." A rational decision had to be made prior to analyzing the data. With respect to the signs that were given ratings by the participants using a scale that ranged from 1 to 5, the authors reasoned that a rating of 3.5 or higher would constitute an acceptable design, and a rating of 4.5 or higher would constitute a highly desirable sign. By using these criteria, the acceptability of a particular design could be evaluated.

Displays on the Major Street (Case I)

The results of the ratings and the rankings of the signs for Case I are shown in Figures 1 and 2 respectively. Frequency distributions for the ratings and rankings are shown in Figures 3 and 4 respectively.

The results clearly show that the basic designs that were simple in nature were preferred over those that displayed a diagram of the area. The design that contained words and color indications (Design 1) and the design that used only color indications (Design 2) to describe the freeway traffic condition were rated relatively high, whereas the designs that included a diagram of the area were rated relatively low. The mean ratings for Designs 1 and 2 were 4.0 and 3.5 respectively, whereas the mean ratings for Designs 3 and 4, which contained diagrams of the area, were 2.5 and 2.4 respectively. On the basis of the preestablished criteria, only Designs 1 and 2 were above the acceptable mean limit.

The data were further analyzed to determine whether there was consistency in the rankings among the participants. Kendall's Coefficient of Concordance, W, was computed for this purpose (4). The coefficient detects the consistency (or lack of consistency) in the ranking of ordinal data. The significance of the coefficient was then tested using the chi-square, χ^2, statistic. The test does not reveal the degree of preference

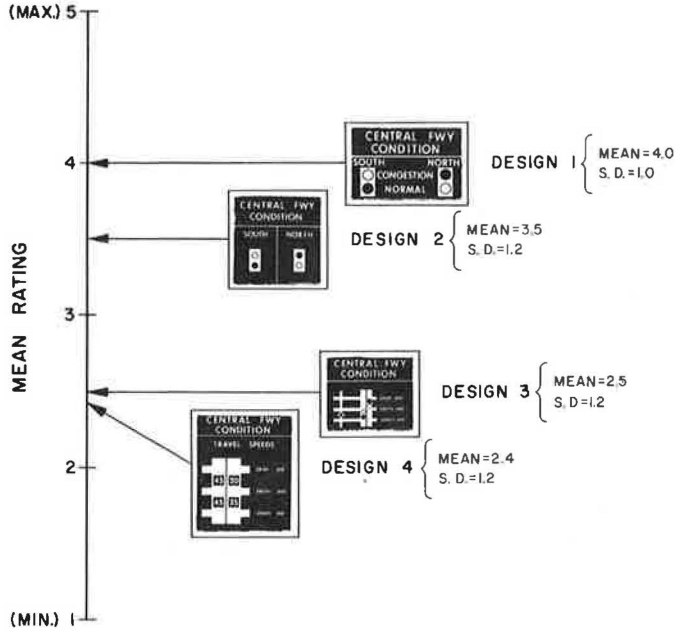
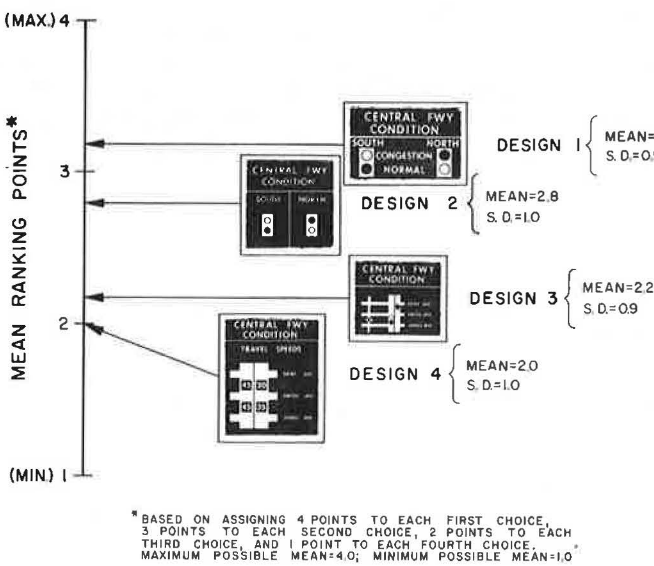


Figure 1. Ratings of sign design alternatives for display on major streets–Case I.

but does determine whether the ranking was consistent among the participants and provides a basis for determining the best estimate of the "true" ranking according to consensus based on the R_j values. The results of the analysis are summarized in Table 2.



*BASED ON ASSIGNING 4 POINTS TO EACH FIRST CHOICE, 3 POINTS TO EACH SECOND CHOICE, 2 POINTS TO EACH THIRD CHOICE, AND 1 POINT TO EACH FOURTH CHOICE. MAXIMUM POSSIBLE MEAN=4.0; MINIMUM POSSIBLE MEAN=1.0

Figure 2. Rankings of sign design alternatives for display on major streets–Case I.

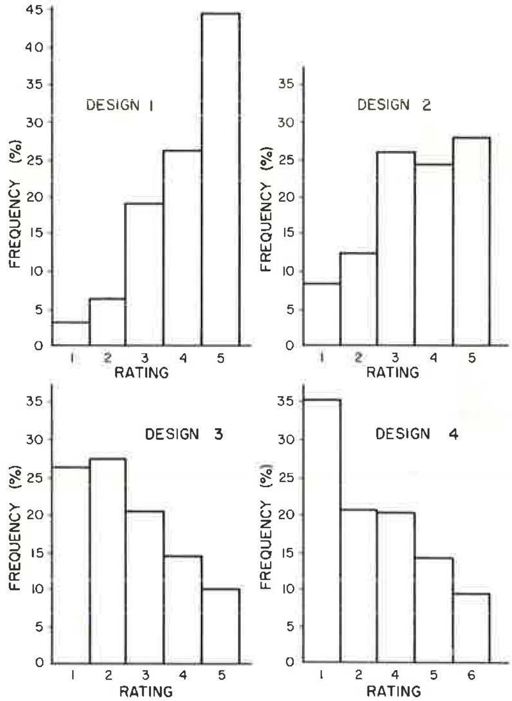


Figure 3. Frequency distributions for ratings of sign design alternatives–Case I.

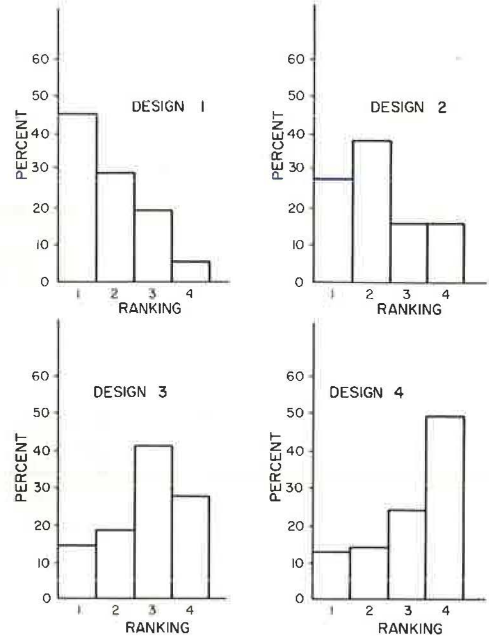


Figure 4. Frequency distributions for rankings of sign design alternatives–Case I.

The results of the test revealed that the coefficient, W, was 0.1841, and the chi-square value of 224.3 was highly significant at the 0.01 level. This means that the respondents had ranked the signs consistently. Based on the values of R_j, the preference for the signs was in the following order: Design 1, Design 2, Design 3, and Design 4.

The results suggest that word messages describing the freeway conditions would be slightly more desirable than a design that was void of qualitative messages. They also reinforce the results of the ratings of each sign. The participants preferred the simple designs over the designs that displayed a diagram of the area.

TABLE 2

SUMMARY OF KENDALL'S TESTS FOR RANKING OF SIGN DESIGN ALTERNATIVES

Alternative	Case I		Case II		Case III	
	R_j Value	Order of Ranking	R_j Value	Order of Ranking	R_j Value	Order of Ranking
Design 1	1,270	1	1,337	1	1,317	1
Design 2	1,127	2	1,049	2	1,220	2
Design 3	890	3	883	3	870	3
Design 4	773	4	861	4	823	4
Kendall Coefficient, W	0.184		0.170		0.206	
Chi Square, χ^2	224.3[a]		210.3[a]		260.8[a]	

[a]Significant at 0.01 level.

Displays at the Entrance Ramps (Case II)

The results of the analysis of the visual displays for use at the entrance ramps are shown in Figures 5 and 6, and frequency distributions are shown in Figures 7 and 8.

It should be noted that one of the basic designs for Case II was slightly different from the pattern listed previously. The second design incorporated color indications to reflect the traffic conditions as well as a diagram of the area to assist the motorist in orienting himself to the facility. In Cases I and III the diagram was not used for this design.

The results again clearly demonstrate that the design that was simple in nature was preferred over the designs that contained a diagram of the area. The mean rating for the design showing a color signal indication and word messages (Design 1) was 3.9, whereas Designs 2, 3, and 4, all of which contained a diagram of the area, had mean ratings of 2.8, 2.7, and 2.4 respectively. Only Design 1 was ranked above the acceptable mean limit of 3.5. The rankings of the alternate signs were consistent with the ratings of the individual signs. The mean ranking for Design 1 was 3.3, whereas the mean rankings for Designs 2, 3, and 4 were 2.6, 2.2, and 2.1 respectively.

The results of Kendall's Coefficient of Concordance for the consistency of ranking of these four signs are given in Table 2. The results again establish that the participants were consistent in the manner in which the signs were ranked. Kendall's coefficient, W, was computed as 0.1697, and the chi-square test was highly significant at the 0.01 level. The tabulated order of ranking for Case II was as follows: Design 1, Design 2, Design 3, and Design 4.

Displays on the Freeway (Case III)

The results of the ratings and the rankings of the sign display alternatives for use on the freeway are shown in Figures 9 and 10. Frequency distributions of the participants' responses to the ratings and rankings are presented in Figures 11 and 12 respectively.

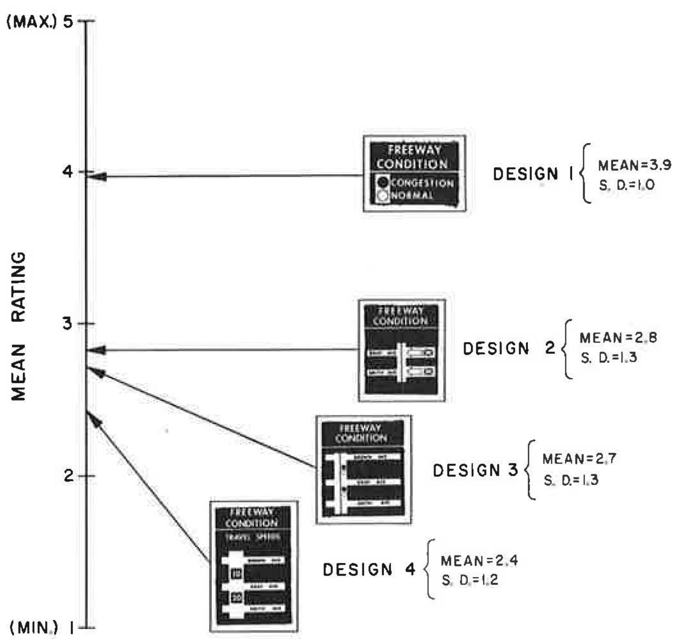
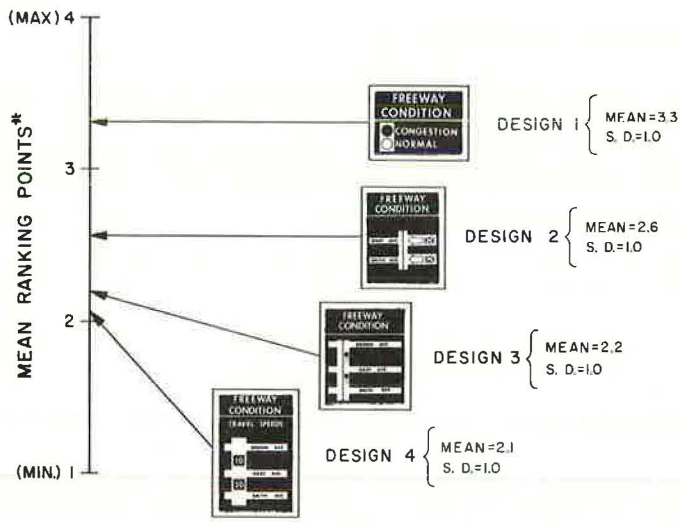


Figure 5. Ratings of sign design alternatives for display at entrance ramps–Case II.



*BASED ON ASSIGNING 4 POINTS TO EACH FIRST CHOICE, 3 POINTS TO EACH SECOND CHOICE, 2 POINTS TO EACH THIRD CHOICE, AND 1 POINT TO EACH FOURTH CHOICE. MAXIMUM POSSIBLE MEAN=4.0; MINIMUM POSSIBLE MEAN=1.0

Figure 6. Rankings of sign design alternatives for display at entrance ramps–Case II.

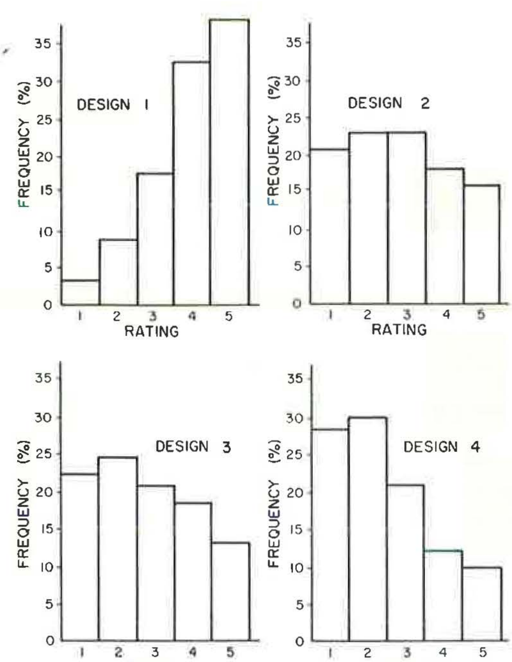


Figure 7. Frequency distributions for ratings of sign design alternatives–Case II.

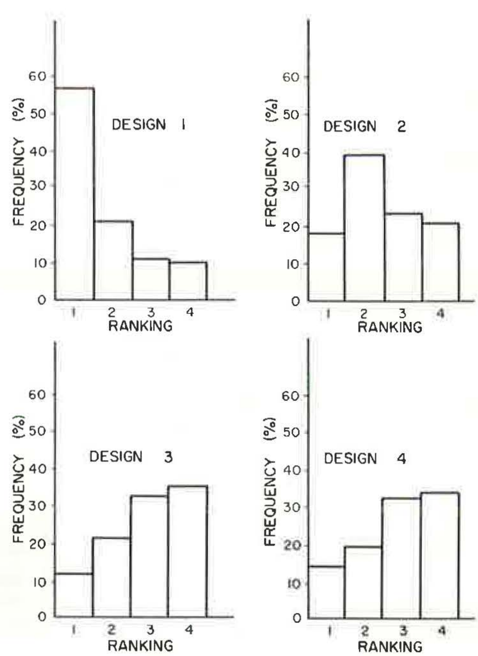


Figure 8. Frequency distributions for rankings of sign design alternatives–Case II.

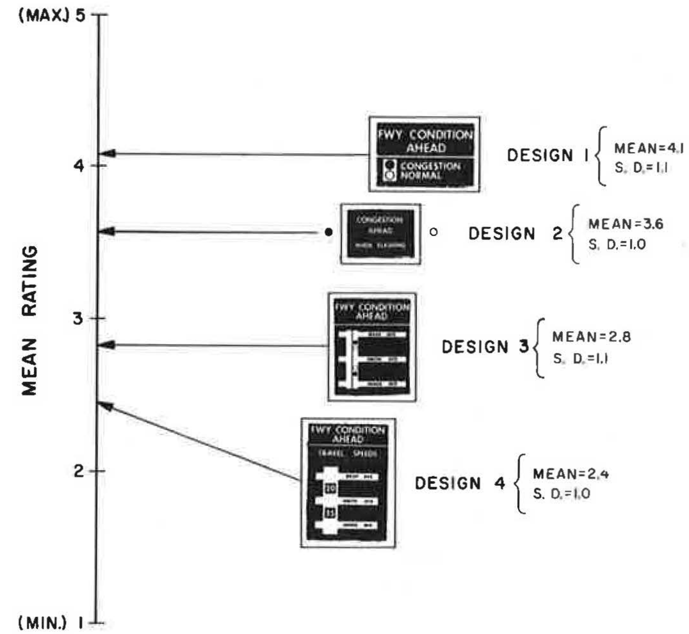
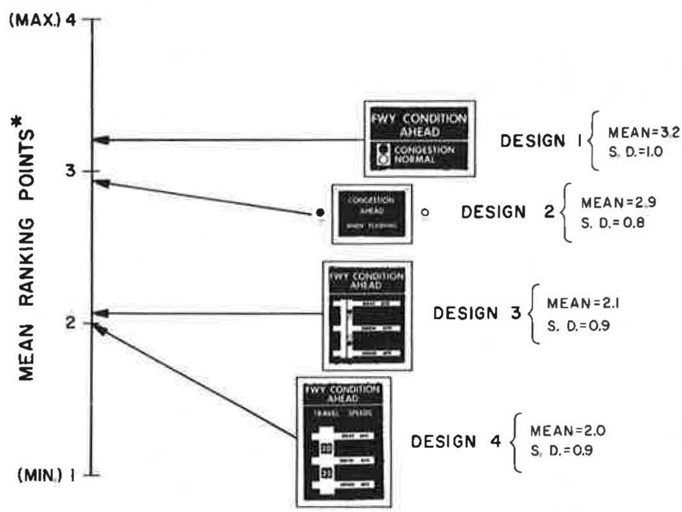


Figure 9. Ratings of sign design alternatives for display on the freeway–Case III.



*BASED ON ASSIGNING 4 POINTS TO EACH FIRST CHOICE, 3 POINTS TO EACH SECOND CHOICE, 2 POINTS TO EACH THIRD CHOICE, AND 1 POINT TO EACH FOURTH CHOICE. MAXIMUM POSSIBLE MEAN=4.0; MINIMUM POSSIBLE MEAN=1.0

Figure 10. Rankings of sign design alternatives for display on the freeway–Case III.

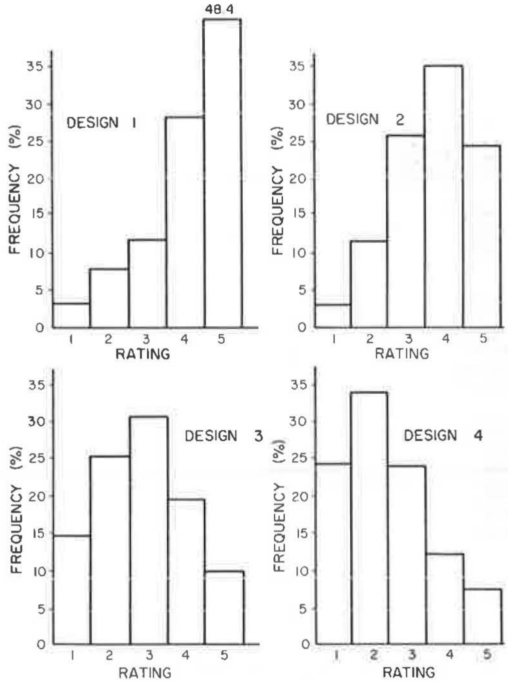


Figure 11. Frequency distributions for ratings of sign design alternatives–Case III.

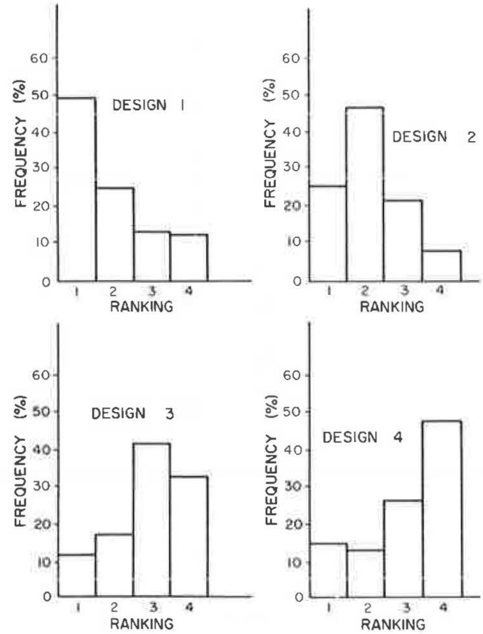


Figure 12. Frequency distributions for rankings of sign design alternatives–Case III.

The results again show a preference for simplicity in design. Designs 1 and 2 had mean ratings of 4.1 and 3.6 respectively, both of which were above the acceptable mean limit. The two designs that incorporated a diagram of the freeway and streets had mean ratings of 2.8 and 2.4. The results of the rankings again show Design 1 to be preferred, followed by Design 2, Design 3, and Design 4 in that order. A summary of Kendall's coefficient test is given in Table 2. The results reveal a consistent pattern in the rankings of the four signs. The ranking of the designs was in the following order: Design 1, Design 2, Design 3, and Design 4.

Summary of Ratings and Rankings of Basic Visual Displays

The mean ratings and rankings for Cases I, II, and III are given in Table 3 to provide a better understanding of the results concerning the design alternatives for a total system. The comparison is made for the purpose of showing the consistency of the four basic designs. There was definite consistency in the ratings and rankings for all three cases.

The results of the ratings and rankings of the four basic designs indicate the preference of the motorist for a simple design. Although it had been conjectured that diagrams providing the driver with an orientation to the freeway and streets would be a valuable asset, the results of the study indicate that this type of display is the least preferred of all the alternatives.

SPECIAL DISPLAYS

A portion of the questionnaire and slide presentation was designed to obtain inputs regarding some special features of visual displays. In one group of questions, the participants were asked to make comparisons among three pairs of signs. Only one

TABLE 3

COMPARISON OF THE PARTICIPANTS' EVALUATION OF VISUAL DISPLAYS FOR ALL THREE HYPOTHETICAL SITUATIONS

Design	Case I		Case II		Case III	
	Mean Rating[a]	Mean Ranking[b]	Mean Rating[a]	Mean Ranking[b]	Mean Rating[a]	Mean Ranking[b]
1	4.0	3.2	3.9	3.3	4.1	3.2
2	3.5	2.8	2.8	2.6	3.6	2.9
3	2.5	2.2	2.7	2.2	2.8	2.1
4	2.4	2.0	2.4	2.1	2.4	2.0

[a]Mean determined by assigning 1 point for rating of 1 (low), 2 points for rating of 2, 3 points for rating of 3, 4 points for rating of 4, and 5 points for rating of 5 (high). Maximum possible mean = 5.0.
[b]Mean determined by assigning 4 points to each first choice, 3 points to each second choice, 2 points to each third choice, and 1 point to each fourth choice. Maximum possible mean = 4.0, minimum possible mean = 1.0.

different feature existed between each pair. The alternatives that were compared are shown in Figure 13. In each of the three comparisons, the participants were asked to indicate their selection from the following choices:

1. Alternative A is best;
2. Alternative B is best; or
3. Alternatives A and B are equally good.

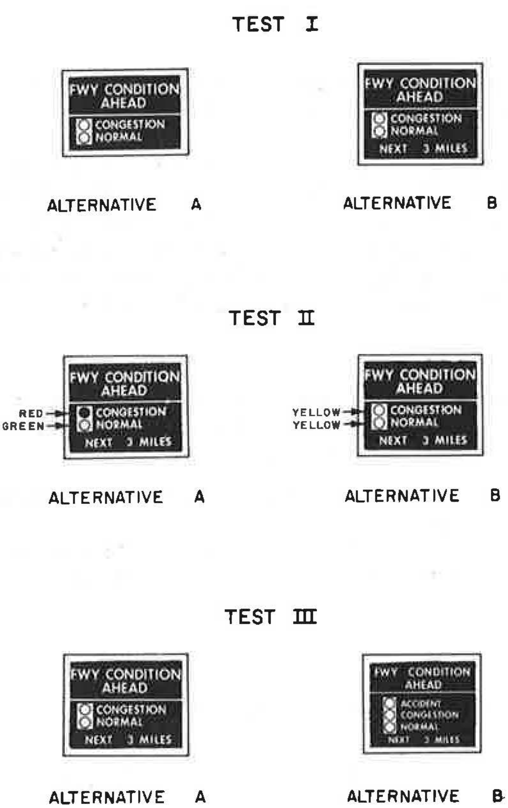


Figure 13. Special feature alternatives for visual display.

Test I was a comparison to determine whether the participants would like to receive information regarding the location of congestion. Test II was used to establish whether the color of the lamps in the visual display would affect the choice of signs. This in essence was one means of measuring the desire for distinct colors to indicate varying degrees of traffic operation. Test III was intended to measure the desire of the motorists to know the nature of the incident that causes the congestion. A summary of the results is given in Table 4.

The results revealed that 87 percent of all respondents preferred information regarding the location of the congested area in addition to the qualitative description of the traffic condition. The results also showed that 7 percent of the respondents were indifferent about receiving the added information concerning the location of the congestion.

TABLE 4

PERCENTAGE OF PREFERENCES FOR SPECIAL VISUAL FEATURES

Test	Alternative A	Alternative B	No Choice
I	6	87	7
II	69	21	10
III	26	57	17

A comparison of the color of the signals in Test II showed that a majority of the respondents preferred the red and green signals in contrast to all yellow ones. A total of 69 percent indicated their preference for the red and green combinations, 21 percent preferred all yellow signals, and 10 percent were indifferent. This result suggests the desirability of color or some other unique characteristic to distinguish the degree of traffic conditions on the freeway.

An analysis of the desirability for knowing the occurrence of an incident in Test III indicated that only slightly more than half of the respondents desired to know that an accident had occurred, in addition to the freeway traffic condition and the length of the congested area: 57 percent of the respondents favored the display that indicated the occurrence of an accident, 26 percent did not desire this added information, and 17 percent were indifferent.

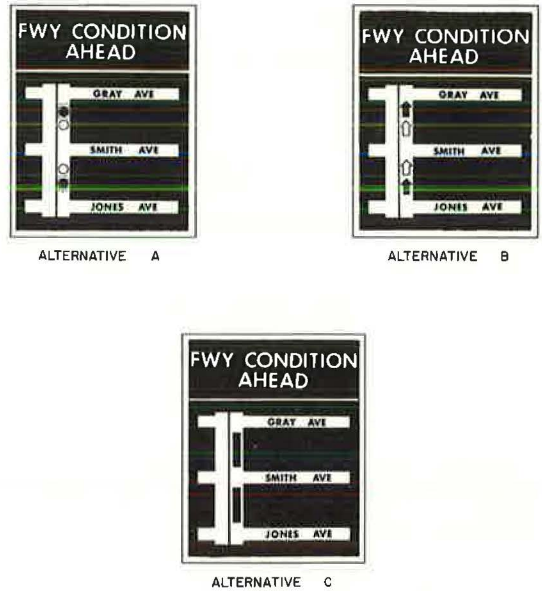


Figure 14. Special symbol alternatives for visual display.

Symbols

It was also of interest to determine the desirability of certain types of symbols that could be used on visual displays. Three alternatives, as shown in Figure 14, were presented to the participants for ranking. The results of this analysis are given in Table 5.

Kendall's Coefficient of Concordance was again employed to determine whether there was a definite degree of consistency of ranking. The coefficient was computed to be 0.0113, and the test of significance revealed a chi-square value of 7.60. The results were not significant at the 0.01 level. The interpretation of the results revealed no meaningful pattern or consistency in the ranking of the three symbols; therefore, there was no reason to believe that an order of preference existed among the symbols.

Color

The participants were asked for their opinions concerning the possible color combinations of a sign giving information about the freeway traffic condition. They were presented the following choices:

1. White letters on a green background, as used for guide signs;
2. Black letters on a yellow background, as used for warning signs;

TABLE 5

PERCENTAGE OF PREFERENCES FOR SYMBOLS ON VISUAL DISPLAYS

Symbol	First Choice	Second Choice	Third Choice	Average Ranking Points[a]	Standard Deviation
Circle	52	25	23	2.3	0.8
Arrow	27	53	20	2.2	0.7
Bar	21	22	57	1.7	0.8

Note: Kendall's Coefficient of Concordance, W = 0.0113; chi-square, χ^2 = 7.60; d.f. = 2.

[a]Based on assigning 3 points for each first choice, 2 points for each second choice, and 3 points for each third choice. Maximum possible mean = 3.0, minimum possible mean = 1.0.

TABLE 6

DRIVER PREFERENCES OF COLORS FOR VISUAL DISPLAYS

Choices	Percentage of Respondents
White letters on green background	22
Black letters on yellow background	9
New color combination	62
No preference	7

3. A new color combination to distinguish these particular signs from all others; or
4. No preference.

Table 6 gives the results, which indicate that the drivers preferred a unique device that clearly distinguishes real-time freeway information from other types of freeway signing.

SUMMARY OF FINDINGS

Based on the analysis of the questionnaire survey administered to 505 licensed drivers, the following findings may be drawn:

1. The respondents preferred real-time freeway information sign displays that were simple in nature. Simple types of displays were consistently preferred over designs containing a diagram that provided the motorists with an orientation to the freeway and streets. Designs 1 and 2 were consistently rated high. Designs 3 and 4 were consistently rated low.
2. A preference was shown for unique design features, such as the use of color, on visual displays to distinguish between usual and abnormal traffic conditions.
3. The respondents indicated a preference for a unique design that distinguishes real-time visual displays from other types of freeway signing.
4. Information with respect to the freeway traffic condition and the location of congestion was preferred over knowledge of only the freeway traffic condition.
5. There was no reason to believe that a preference existed for any of the following symbols that could be used on a real-time visual display: circle, arrow, or bar.

REFERENCES

1. McCasland, W.R. Freeway Ramp Control System. Texas Transportation Institute Research Report 24-26, Aug. 1969.
2. Dudek, C.L., and Cummings, D. Application of Commercial Radio to Freeway Communications—A Study of Driver Attitudes. Texas Transportation Institute Research Report 139-3, July 1970.
3. Dudek, C.L., and Jones, H.B. Real-Time Information Needs for Urban Freeway Drivers. Texas Transportation Institute Research Report 139-4, Aug. 1970.
4. Siegel, S. Nonparametric Statistics. McGraw-Hill Book Co., New York, 1956.

DISCUSSION

Marshall Jacks, Federal Highway Administration

The inventory of Freeway Surveillance and Operational Control Activities prepared by the Committee on Freeway Operations and issued in June 1970 as Highway Research Circular 108 indicates that 25 agencies in 16 states reported some form of freeway operational and control activity. Sixteen of the reporting agencies indicated that changeable message signs are being used to provide travel information or variable speed limits. In all probability an inventory of the same activities taken 3 years from now would show an increase of 100 percent or more in actual mileage of freeways utilizing similar concepts.

The continuing use of surveillance and control concepts for improvement of urban freeway operation dictates the necessity for the development of standards for the application of real-time visual information displays. To be effective these displays must inform the motorist in a manner most easily comprehended by him. The research that this paper describes is directed toward a subject that can be immediately applied, a factor of critical importance.

The evaluation procedure was apparently well designed to accomplish the goals effectively. The various design alternatives were such that the subject drivers were asked to indicate preference or lack of preference on the basis of conceptual differences in information forms. The findings seem to indicate clearly that the top preferences were for (a) simplicity of design; (b) use of combination word and color symbol messages; (c) use of red and green color symbols to denote traffic congestion and normal movement respectively; and (d) use of a unique design for the real-time visual display device. Additional preferences of lesser priority were indicated for information regarding the location of congestion and whether or not it is caused by an accident. Diagrammatic displays indicating congestion or travel speeds in certain areas and preferences for a specific symbol received relatively low priority.

The findings indicate that drivers prefer display concepts similar to those that have evolved over a period of time into standards. These concepts, which are now standard, include three-color vehicular traffic signals, two-color pedestrian signals utilizing word messages, and unique shapes or other design features for various signs. Designs of these standard devices have evolved over periods of time during which features of the individual displays have undergone various changes. However, we are now at a stage of development of standards where certain shapes and colors have been reserved for specific uses. In that connection it is suggested that, because there was no clear-cut preference for circular or arrow symbols over bars, future efforts should consider the use of red and green bars or other unique shapes rather than circular or arrow symbols.

As a general statement, it is my opinion that this report describes the results of an effective research effort. It represents an early stage in what should be a multistage research effort to develop effective real-time visual displays for urban freeways. The proper follow-up of this study can accelerate the development of standards for these devices. To promote this accomplishment the following suggestions are offered:

1. Similar evaluations using the identical or slightly revised designs should be made by using subjects from different geographical areas.
2. If the findings of this work are correlated in previous or future studies, designs of low priority should be dropped from consideration and prototypes of the high-priority designs developed and tested for driver reaction.
3. Following successful testing, recommendations for development of standards should be made.

As indicated earlier, this subject is one for which there is an opportunity for immediate practical application. The proper response from the research community can accelerate the transition from basic research to development of application and design standards.

F. Lehman, Newark College of Engineering

The authors have made a notable contribution in this paper by pointing out the importance of driver involvement in and acceptance of traffic control devices. Their conclusions seem quite consistent with driver response to other types of highway signs.

Driver acceptance and/or credibility is of course a major concern in the decision to spend public funds for real-time information displays. From experience in California and New Jersey, there is apparently a large credibility gap in real-time speed control displays used on freeways. Unfortunately, this situation has a long history founded on unreasonably low posted speed limits in many small towns.

In New Jersey real-time synchronized signal-ahead signs have met with good acceptance. One reason, no doubt, is that they have been judiciously placed where the driver needs the information for proper control. Although no quantitative measurements have been made, there is the general belief that the potential for rear-end collisions has been reduced and traffic flow has been improved.

Although the authors state that the designs used were similar to those in use in other cities, it would seem to be in order to state what the design requirements should be and then to compare driver preferences with how the given designs fit the design requirements. Essentially, the requirements are the five given in the Manual on Uniform Traffic Control Devices. Implied in the requirement of commanding attention is that of ease of differentiation among other points of attention.

From the general requirement of ease of comprehension it is quite apparent that Designs 1 and 2 are superior to Designs 3 and 4. There is more information conveyed in Designs 3 and 4, but this information is more difficult to sort out from the diagrams. This is particularly true on first presentation. Greater familiarity with this latter type of sign may well improve its rating by the driver.

It was tacitly assumed by the authors that a design is "good" if a majority of drivers perceive it as such. Yet according to the philosophy of design applied to roadway elements it is usually assumed that the design will normally accommodate all drivers. This suggests the criterion that the design should be operationally satisfactory for all drivers under normal driving conditions. This would mean that no driver would give a poor rating to an acceptable design. In the sample used by the authors about 3 percent of the drivers gave a poor rating to Design 1, which was judged best by the majority. Considering, however, the particular purpose of this type of sign, the low percentage is not very important.

On the subject of testing procedures, several questions might well be raised. The main one is how closely the test situation approximates actual driver behavior. In terms of conclusions, would the differences between the real situation and the simulated conditions be expected to change appreciably the test findings? Out of these come other questions: How does the participant's viewing time during the slide presentation compare with viewing time in the driving situation? How does the apparent size in the slide presentation compare with the apparent size of the real display as viewed by the driver? Again, greater familiarity with the displays used could possibly change the participant's preferences.

A final question is concerned with how to determine what constitutes "good" from the participant's response. The authors themselves made that judgment by selecting a lower bound of 3.5 in their rating scale of 1 to 5. Would it not be simpler and more meaningful to have participants make the judgment by allowing them to select from a four- or five-part attribute scale such as very good, good, fair, poor? Then assigned numbers can be applied later to the selected attributes for the purpose of obtaining an average rating.

AUTHORS' CLOSURE

The authors appreciate the reviews by Marshall Jacks and Fred Lehman.

Jacks has pointed out the analogy of the development of real-time information displays to the highway display concepts that have evolved over the years into present standards. He has challenged the research community to accelerate the transition of real-time information systems development from basic research to an operational reality and has offered sound suggestions toward this goal. As part of our continuing research program, prototype changeable message signs will be installed in the Gulf Freeway corridor to complement the existing computer-controlled facilities. Field studies will then be conducted to evaluate their effectiveness under the real situation. Similar programs are under way in other parts of the country.

Lehman has raised several important questions relating to the acceptance criterion and the testing procedures. An acceptance criterion was established so that the features of each design that are acceptable to the motorists could be evaluated. These features will then be incorporated into the prototype designs for further evaluation in the field.

Because the visual displays were projected on a screen, the physical size of the displays was consequently smaller than the signs that will be erected on the highway. However, because of the close proximity of the participants to the screen, the relative sizes of the projected displays were approximately the same as typical changeable message signs as viewed from the automobile.

The participants were required to look at and evaluate each display within 10 seconds. The actual observation time of the display would, therefore, compare favorably with the expected viewing time in the driving situation.

A numerical scale ranging from 1 to 5 was used in the questionnaire to rate the various sign designs. A "low" was affixed to the number 1 and a "high" to the number 5 to give direction to the scale. This method was used for two reasons: In the first place, this type of scale does not require that all subjects agree on the definition of a stated attribute. For example, the word "fair" may be considered to be a positive response by some, whereas others may consider it negative. This problem is somewhat negated when a numerical scale is used. Secondly, the results of a questionnaire survey using the numerical scale are easily quantified, and group scores may be directly compared to a given individual's response without transcribing or otherwise converting the scale.

TRAFFIC DELAY AND WARRANTS FOR CONTROL DEVICES

Walter C. Vodrazka, Wisconsin State University—Platteville;
Clyde E. Lee, University of Texas at Austin; and
Herman E. Haenel, Texas Highway Department

Delay to vehicular and pedestrian traffic is one of several criteria frequently used in the selection and evaluation of traffic control devices. In the past, there has been no practical technique for measuring, recording, and analyzing delay data in sufficient quantities to provide a sound basis either for developing delay-based warrants or for determining the relative effectiveness of various control devices in limiting delay to tolerable values. A unique digital delay data recorder that was developed and used successfully for collecting over 240 hours of field data at 19 different intersections is described. This device records up to 12 channels of information from human observers or traffic signal controllers in a form directly suitable for computer processing. Complete summary statistics, which can involve processing as many as 360,000 data items, can be obtained on an overnight basis for 6 hours of field observation. These statistics include 19 delay-related traffic parameters and may be summarized for any selected time intervals. Analysis of delay data from field studies conducted over a 3-year period facilitated the development of a new set of minimum volume warrants for the installation of four-way stop-sign control and the validation of a proposed set of traffic volume warrants for the installation of actuated signal control. These warrants are presented in a tabular and graphic form that is suitable for ready application by practicing traffic engineers.

•TRAFFIC control devices ranging in complexity from signs and pavement markings to sophisticated signal systems are used to assign the right-of-way alternately to traffic on the several approaches to street and highway intersections. The relative effectiveness of these devices can be measured in terms of delay to motorists and number of accidents. The objective is to move the maximum volume of traffic safely through the intersection with minimum delay.

Even though delay to vehicular and pedestrian traffic is a frequently used criterion for selecting and evaluating traffic control devices, no practical technique for measuring, recording, and analyzing delay data in sufficient quantities to serve as the basis for delay-based warrants has been available. A unique digital delay data recorder that was used successfully over a 3-year period for collecting some 240 hours of field data at 19 different intersections is described in this paper. Analysis of these data facilitated the development of a set of minimum volume warrants for installing four-way stops and provided a basis for validating a set of proposed traffic volume warrants for actuated signals.

Several potential applications for the delay recording equipment and the associated analysis techniques are suggested. These include collecting quantitative information for before-and-after studies of traffic control efficiency.

D3 RECORDER

In order to record the large amounts of data needed for studying vehicular delay characteristics at intersections in a form directly suitable for computer processing, the dig-

Sponsored by Committee on Traffic Control Devices.

ital delay data recorder (D3 Recorder) was developed. This electromechanical instrument is capable of recording automatically coded switch closures that indicate the number of stopped vehicles, the cumulative traffic volume, and the signal indication for each approach lane (up to 12) on a moment-by-moment basis for extended periods of several hours.

Although bulky in its present form, the equipment is easily transported and can be set up at a field site in about 30 min. From 6 to 18 observers are required to provide instantaneous input information regarding the number of stopped vehicles and the number of vehicles that enter the intersection. The observer accomplishes this simply by actuating a push-button switch. The signal indication on each approach is sensed by a direct wire connection to the appropriate power contact at the signal controller.

Data are converted to a digital format and punched directly onto paper tape at the study site. The punched paper tape serves as an inexpensive intermediate storage location inasmuch as the data must be transferred finally to magnetic tape for computer processing. Most computation centers, however, have facilities for reading punched paper tape on a routine basis. Other advantages of the punched-paper-tape recording are that an experienced operator can spot-check the data by visual inspection in the field and that the record is permanent.

Field experience with the recorder over a 3-year period proved it to be quite simple and efficient to operate. A full-time crew of eight men, mostly high-school students, were trained for studies that were made during the summer months of 1966 and 1967. Less than 30 min of instruction was required for each new observer. A review of data that were input by duplicate observers in a special study indicated that reliable information could be obtained even with this minimal observer training.

Each observer was equipped with a data input module containing two push-button switches (add and subtract) and a remote indicating counter. This module was connected to the recorder by a small multiconductor electrical cable 200 ft long so that the observer could be as inconspicuous as possible while watching the traffic. All that was required of the delay observers was to keep an instantaneous count of the number of stopped vehicles showing on the indicating counter. The traffic volume observers simply added a count for each vehicle that entered and cleared the intersection.

The recorder was programmed to scan sequentially all input channels and record channel identification, signal indication, number of stopped vehicles, and number of vehicles that had entered the intersection. When 12 channels were being used, each channel was scanned every 3 sec; when only six channels were used, the scan rate was once each 1.44 sec. These rates were selected as a suitable compromise when considering equipment complexity, quantity of data to be processed, driver and observer reaction time, and statistical sampling.

From the data recorded in the field, several pertinent values related to delay were calculated. Delay relationships for each individual approach and for the intersection as a whole were summarized for several selected time periods. After data from several studies were evaluated, a 15-min analysis period was chosen for continued use. The values that were calculated for each approach for 15-min periods were

1. Traffic volume;
2. Total vehicle-seconds of delay;
3. Total number of vehicles stopped;
4. Average delay per vehicle;
5. Average delay per vehicle stopped;
6. Percent of vehicles stopped;
7. Total green time;
8. Number of complete cycles;
9. Average green time per cycle; and
10. Average cycle length.

The first six items were calculated for the sum of all approaches as well. The seventh and ninth items were characteristic of a given direction, and the eighth and tenth were characteristic of the intersection control. Attempts were made to calculate other relationships such as the vehicle-seconds of delay due to left turns, the total number of

TYPICAL CALCULATIONS FOR A 15-MINUTE PERIOD

WOODROW AND KOENIG JULY 25, 1967 0715 TO 0915 FULL ACTUATED

TIME PERIOD 800 - 815

COMPUTED INFORMATION	APPROACH A	APPROACH B	APPROACH C	APPROACH D
TRAFFIC VOLUME	97.00	77.00	28.00	84.00
TOTAL VEH-SECS OF DELAY	578.88	544.32	220.32	612.00
VEH-SECS OF DELAY DUE TO LEFT TURNS	4.32	7.20	18.72	8.64
TOTAL NO OF VEHS STOPPED	40.00	49.00	14.00	44.00
TOTAL NO OF STOPS	43.00	51.00	18.00	47.00
AVERAGE DELAY PER VEHICLE STOPPED	14.47	11.11	15.74	13.91
AVERAGE DELAY PER VEHICLE	5.97	7.07	7.87	7.29
AVERAGE DELAY TO THE FIRST VEHICLE	15.16	14.47	14.94	14.23
PERCENTAGE OF VEHICLES STOPPED	41.24	63.64	50.00	52.38
TOTAL GREEN TIME	406.08	485.28	410.40	485.28
AVERAGE GREEN TIME PER CYCLE	17.10	19.56	17.16	19.56
NUMBER OF CYCLES	24.00	24.00	24.00	24.00
AVERAGE LENGTH OF CYCLE	36.96	36.84	36.90	36.84
TOTAL X TIME	0.	0.	0.	0.

TOTAL TRAFFIC VOLUME	286.00
TOTAL VEH-SECS OF DELAY ALL APP.	1955.52
TOTAL NUMBER OF VEHS STOPPED ALL APP.	147.00
TOTAL NUMBER OF STOPS ALL APP.	159.00
AVERAGE DELAY PER VEH STOPPED ALL APP.	13.30
AVERAGE DELAY PER VEHICLE ALL APP.	6.84
PERCENTAGE OF VEHICLES STOPPED ALL APP.	51.40
AVERAGE DELAY TO FIRST VEH ALL APP.	14.65

APPROACH A IS SOUTHBOUND
APPROACH B IS WESTBOUND
APPROACH C IS NORTHBOUND
APPROACH D IS EASTBOUND

Figure 1. Example of calculations.

stops, and the average delay to the first vehicle. There were, however, difficulties in calculating these values, which limited their usefulness in the analysis of intersection delay characteristics.

An example of the calculations made for a 15-min period at a typical intersection is shown in Figure 1. The 10 values previously listed are given along with one additional value, "Total X Time." This value refers to the total time in the time period during which data were missing or otherwise unusable.

The traffic volume was determined as the difference between the recorded volumes at the beginning and end of the time period under study.

Vehicle-seconds of delay were computed as the product of the sum of the indicated number of stopped vehicles for each recording interval in the time period and the length of the interval, which was either 1.44 or 3.00 sec.

When the indicated number of stopped vehicles is plotted as the ordinate versus the midpoint of each recording interval on a continuous time scale, the area under the curve is equivalent to this calculation.

The total number of vehicles stopped was determined for each approach by counting the increases in the indicated number of stopped vehicles during each red signal indication and in the first few seconds of green signal time. Here, the assumption was that an arriving vehicle was forced to stop at the rear of the queue. When an increase in the indicated number of stopped vehicles occurred during the green signal indication, it was observed in the field that this was most often due to a previously stopped vehicle waiting to make a left turn.

By addding both types of increases we get a quantity called the total number of stops. If an increase in the indicated number of stopped vehicles occurred during a green signal, the number of stopped vehicles was accumulated for each interval until a decrease was observed. The vehicle-seconds of delay due to left turns were then calculated by multiplying this accumulated number by the recording interval length.

Of course, this method of determining the number of vehicles stopped and the left-turn delay is not foolproof. A vehicle could arrive at the rear of a queue just as a vehicle departs from the front, and the indicated number of stopped vehicles would remain unchanged.

The average delay per vehicle and per vehicle stopped was calculated by dividing the total vehicle-seconds of delay by the volume and the total number of vehicles stopped respectively.

Total green time was measured by counting the number of intervals in which the green signal was displayed and multiplying by the interval length. The determination of the number of complete signal cycles was slightly more complicated. The interval at which the red signal indication first changed to green was noted. The next time that red changed to green marked the end of the first cycle. Thus, the total number of times that red changed to green during the time period under study was one more than the number of complete cycles. The average green time per complete cycle and the average length of a complete cycle were then easily computed.

The average delay to the first vehicle was calculated as the total delay to all first vehicles observed in the time period divided by the number of first vehicles observed. A first vehicle was considered observed at the first recording interval that indicated at least one stopped vehicle after the preceding recording interval had indicated no stopped vehicles, subject to the limitation that only those events taking place during a red signal indication would be counted. For each first vehicle observed, the number of recording intervals was counted, up to but not including the interval when the indicated number of stopped vehicles decreased. The total of these intervals multiplied by the interval length yielded the total delay to first vehicles. A precaution was taken so that once a first vehicle was observed, the associated decrease had to occur in the same time period. Otherwise, the observation was counted for the next time period. Those values that were applicable to the intersection as a whole were obtained by appropriate summation and subsequent calculation.

There are obvious advantages to having detailed quantitative information concerning traffic performance at an intersection. The D3 Recorder represents a first-generation attempt at providing a practical tool for obtaining such information. Even though the feasibility of developing workable equipment and analysis techniques has been demonstrated, there are some disadvantages to be considered. The principal disadvantage associated with the use of the recorder is the large number of observers required. Drivers are curious about the presence of the equipment, the people, and the wires connected to the signal controller, and this curiosity tends to result in a gaper's block in the traffic stream. A sign reading "Traffic Survey" placed on the recorder was quite effective in minimizing this phenomenon. The other major disadvantage in the recorder was the result of its being constructed before sophisticated electronic switching components were readily available. The recorder can now be designed to fit into a small suitcase and weigh less than 50 lb by using solid-state devices that are commercially available.

STUDY SITES

All of the sites selected for study were located in Austin, Texas, except for one intersection in San Antonio and a special before-and-after study of a diamond interchange on the Gulf Freeway in Houston. A total of 19 intersections were selected at which 124 individual studies, consisting of approximately 240 hours of observed data, were performed.

Except for the diamond interchange, all the intersections had four approaches and were essentially right-angle crossings. The sites were generally situated in suburban areas that were classified as either outlying business districts or residential fringe areas. Parking was prohibited on all approaches in virtually all instances, and sight distances were generally adequate. The volume of pedestrian and truck traffic at each intersection location was negligible.

Every effort was made to select intersections that had similar geometric proportions and that included several in each control type category. It was also desired that each intersection be isolated so that the delay characteristics of the type of control were measured without being greatly influenced by nearby similarly controlled intersections; however, this was virtually impossible to do.

The number of two-way and four-way stop-sign controlled intersections in the vicinity of Austin that had appreciable traffic volumes was limited. Thus, the stop-controlled intersections included in this study cover a range of geometrical proportions and cannot be classified by a simple set of characteristics.

Only one pretimed intersection that was deemed suitable for inclusion in this research effort was found in Austin. Not one semi-actuated controlled intersection was found suitable.

However, several similar full-actuated intersections were studied. These intersections were first studied in their existing condition and then the controller settings were changed. The initial, maximum, and vehicle intervals were varied for separate studies. The controller was then made to operate as a pretimed and a semi-actuated controller, and more delay data were recorded. This made it possible to investigate the delay characteristics at an intersection operating under several different control modes.

STOP-SIGN CONTROLLED INTERSECTIONS

Three intersections in Austin (29th and Jefferson, 19th and Chicon, and 38th and Speedway) were studied under two-way stop control at various times during the day, including the morning and evening peak periods as well as a midday period. Preliminary work showed no discernible evidence that delay characteristics were affected by the time of day for the data recorded in this study.

In Figure 2, the sum of vehicle delay on the two stop-sign controlled approaches is plotted as a function of the total volume on all four approaches for 15-min intervals. It may be observed in this figure that delay increases rather gradually to a volume of about 200 to 250 vehicles per 15-min interval. At this volume, a break in the curve occurs and delay increases quite sharply with further volume increases.

Five intersections in Austin (Woodrow and Justin, North Loop and Woodrow, 19th and Chicon, 15th and Congress, and Balcones and Hancock) were studied under four-way stop-sign control, each at various times during the day.

The relationship between total delay and total volume for four-way stop control is shown in Figure 3. A direct comparison of Figures 2 and 3 illustrates the larger total delay experienced at four-way stop controlled intersections.

For a given volume on an approach, the total delay and the average delay were greatly reduced for a stop-sign controlled approach when intersection control was changed from two-way to four-way stop control. However, the total delay experienced on all intersection approaches is greater for four-way than for two-way stop control for equal intersection volumes. This, of course, is because all vehicles must stop and suffer delay under four-way stop control. Thus, a reduction in average delay experience (for the stopped vehicles) must be traded off with an increase in total delay when converting from two-way to four-way stop control.

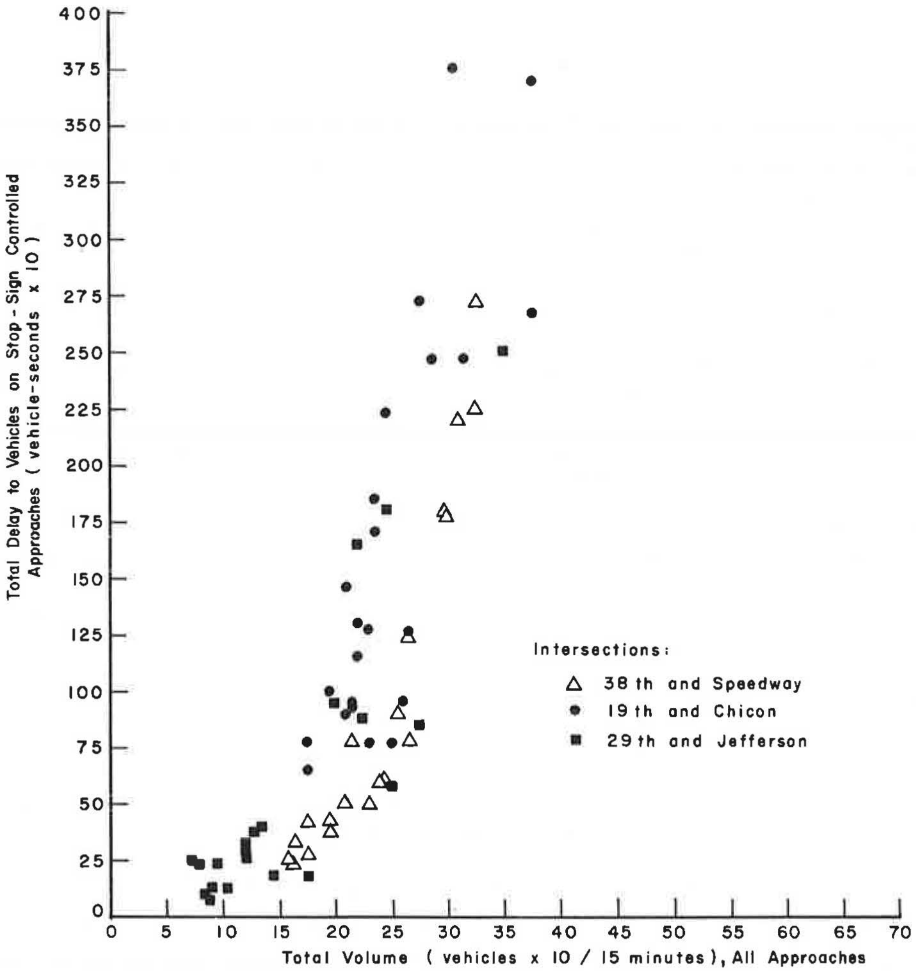


Figure 2. Total delay to vehicles on stop-sign controlled approaches versus total volume (two-way stop control, 15-min intervals).

It is important to note that in Figure 3 the plotted data were observed at five different intersections. The consistency of these data is rather marked and indicates that a strong relationship exists. A square-root transformation was made on the delay variable. A regression yielded a relationship having an R^2 of 0.984 with the following functional form:

$$y = (18.95 + 0.00044\,x^2)^2$$

where

y = the total vehicle-seconds of delay per 15-min interval and
x = the total vehicular volume per 15-min interval.

This relationship is plotted in Figure 3. A square-root transformation is often of value in working with data that are Poisson-distributed. The hypothesis of Poisson-distributed data, however, could not easily be tested.

In comparing two-way and four-way stop operation, reference to Figures 1 and 2 shows that total delay began to increase very rapidly at total volumes of from 200 to 250 vehicles per 15-min interval for two-way stops and from 250 to 300 vehicles per 15-min

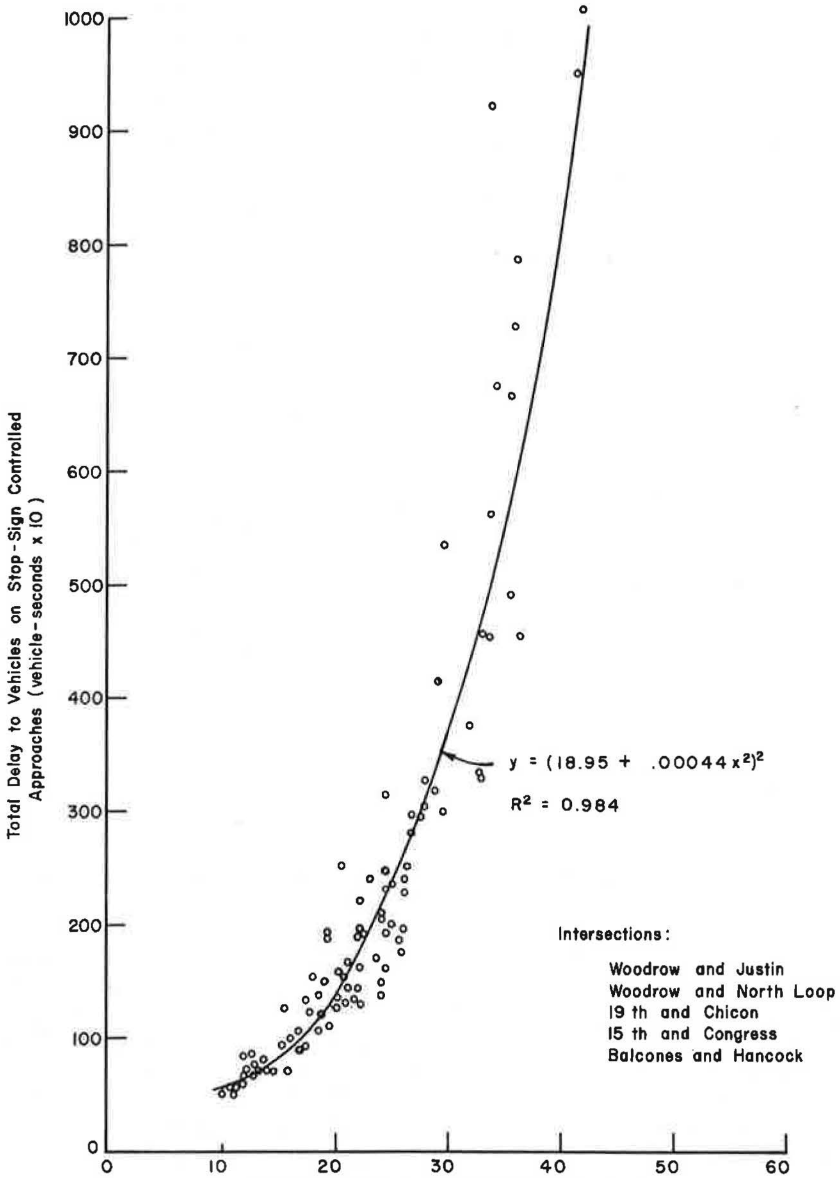


Figure 3. Total delay versus total volume (five intersections, four-way stop control, 15-min intervals).

interval for four-way stops. These 15-min volumes of 250 and 300 vehicles may be termed the critical volumes for two-way and four-way stops respectively. The corresponding critical hourly volumes are 750 and 900 as determined from analysis of the same field data on an hourly basis.

In studying the characteristics of intersections, many variables deserve consideration, including directional volumes, turning movements, approach speeds, width and number of lanes, truck and pedestrian traffic, intersection geometry, and distance to adjacent intersections, among others. In almost all cases in this study, such factors as directional volumes, lane widths, intersection geometry, and the location of adjacent intersections were measured or could be determined. Truck and pedestrian traffic was very minor and was considered to have negligible effects in most instances.

Few data on turning movements and approach speeds were available. Some manual counts of left-turn movements were kept, but these did not appear to have much influence on the delay characteristics of the intersections studied. In general, for almost all variables other than delay and volume, the range of the recorded variable was so limited that its significance, if any, was masked. The geometric layout of each intersection studied under stop-sign control will show differences in geometry, but these seemingly did not influence delay characteristics.

It is of particular importance to recognize that no conclusion is drawn regarding the irrelevance of these variables to delay characteristics other than in the limited range to which the variables were confined in the studies reported. Additional studies designed especially to measure the influence of these variables must be carried out if the variables are to be understood thoroughly.

WARRANTS FOR FOUR-WAY STOP

The generally accepted warrants pertaining to the installation of stop signs, yield signs, and the various types of signals are published in the Manual on Uniform Traffic Control Devices (1). The purpose of these signs and signals is to assign right-of-way to traffic on the approaches to an intersection where conditions of hazard exist such that uncontrolled intersection operation is not feasible and the normal rule, "the vehicle on the right has the right-of-way," cannot be applied safely or efficiently.

The normal hierarchy of control devices, with respect to both cost and effectiveness, is probably the following: yield sign; two-way stop sign; four-way stop sign; and the several signal configurations, including pretimed, semi-actuated, full-actuated, and volume-density devices.

In general, a yield sign is employed for special intersection configurations such as channelized right-turn lanes, intersections with a divided highway, or ramp entrances with inadequate or no acceleration lanes. Yield signs should also be considered applicable at intersections where stop signs are warranted but visibility and speed conditions are such that a full stop is not necessary for safety.

Stop signs may be warranted at almost any intersection of a minor road with a main road or an intersection of two main roads, at an unsignalized intersection in a signalized area, and at railroad crossings. However, stop signs are warranted at any intersection where hazard or accident history indicates a need for stop-sign control. Generally, the two opposing minor-stream flows are stopped while the larger, major-stream flows are not stopped. Under certain conditions, all four approach flows must stop. This necessitates four-way stop control, for which the Manual (1) lists more specific warrants, as opposed to the general policy outlined for yield and two-way stop control.

A four-way stop may be used as a temporary measure at an intersection to be signalized and at an intersection with turning and right-angle accidents accumulating to at least five within a 12-month period. In addition, certain minimum traffic volumes are established:

1. The total, all-approach vehicular volume must average at least 500 vehicles per hour for any 8 hours of an average day.
2. The combined vehicular and pedestrian volume from the minor approaches must average at least 200 units per hour for the same 8 hours with an average delay of 30 sec per vehicle or more for the minor-street traffic during the maximum hour.
3. The volume warrants are reduced to 70 percent of those previously given when the 85th percentile approach speed of major-street traffic exceeds 40 mph.

The Manual suggests, among several qualifications regarding the installation of stop signs, that a four-way stop not be used where the traffic volumes on the intersecting streets are very unequal. If the volumes are heavy enough to warrant additional controls in this instance, a signal installation might be preferable.

The results of the study reported here show that the total delay experienced at four-way stop intersections is virtually unaffected by traffic splits ranging from 50/50 to about

80/20 (Fig. 3) at total intersection volumes up to about 1,400 vehicles per hour for 4 by 4 intersections and up to about 1,100 vehicles per hour for 2 by 2 intersections. Furthermore, the data give no indication of any influence on delay due to the traffic split when analyzed on an approach basis. These were the greatest hourly volumes observed at these intersection types in this study. Of course, this does not imply that the delay experience at these volumes is satisfactory.

Therefore, it is suggested that when the installation of a four-way stop sign is under consideration, the traffic split should not be a factor in making the decision. At larger volumes at which the traffic split might be a factor, a signal installation, rather than a four-way stop installation, should be given consideration.

The results of this study show that the total delay is greater at four-way stops than at two-way stops for a given total volume throughout the range of total volumes observed. Thus, a warrant for four-way stops should be designed to limit the average delay experience rather than the total delay experience.

The warrants given in the Manual set two main conditions: first, to impose a minimum average volume over an 8-hour period and, second, to impose a minimum deviation from the maximum-hour volume such that an average delay to stopped vehicles of at least 30 sec is experienced during the maximum hour. This means that at least 4, and possibly 5 or 6, of the 8 hours will have volumes under 500 vehicles per hour, but the highest hour must have between 900 and 1,000 vehicles per hour.

It would seem more realistic to set a limit on average delay and to work backwards to establish a set of volume warrants. The numbers of hours to be used in computing an average volume must be selected first. As stated previously, 4 to 6 of the 8 hours would have volumes under 500 vehicles per hour, which is below the critical volume of 750 vehicles per hour as mentioned previously in this paper for two-way stops. In establishing a new warrant, it was decided to use 4 hours. Both of the 2-hour periods would probably center around each of the morning and afternoon peak periods.

The following procedure was used in establishing the warrants:

1. A range for tolerable average delay was selected. Analysis of the data shown in Figure 2 yields average delays to stopped vehicles of 20, 30, and 35 sec per vehicle for 15-min total intersection volumes of 220, 285, and 320 respectively. These average delays are characteristic of through-to-stopped-vehicle ratios of about 80/20 to 60/40. Average delays are lower for ratios outside this range.

2. A range for peak-hour factor was selected. A peak-hour factor was necessary to convert the 15-min volume of step 1 to a maximum-hour volume. Three ranges of peak-hour factors were used—0.75 to 0.80, 0.80 to 0.85, and 0.85 to 0.90:

$$\text{PHF} = \frac{\text{Peak-Hour Volume}}{4\ (\text{Peak 15-Min Volume})}$$

3. A peak-period factor was selected. This factor was used to convert the maximum-hour volume of step 2 to the average hourly volume observed during the 2-hour peak period. The peak-period factor is similar to the well-known peak-hour factor and is calculated in the following manner:

$$\text{PPF} = \frac{\text{Sum of Volumes for Four Peak Hours}}{\text{4 times Maximum-Hour Volume}}$$

or

$$\text{PPF} = \frac{\text{Average Hourly Volume}}{\text{Maximum-Hour Volume}}$$

Thus, the average hourly volume for the 4-hour period is the product of the maximum-hour volume and the peak-period factor. Four peak-period factors that were representative of the observed data from this study were used in this analysis: 0.60, 0.70, 0.80, and 0.90.

The application of this procedure resulted in the establishment of the minimum volume warrants for four-way stop signs (Table 1). It is the province of the engineer in charge to decide on the average delay and peaking factors to be used in each specific case. However, it is recommended that (a) the peaking factors be based on field observations (or local experience), (b) an average delay of 30 sec per stopped vehicle be used, and (c) the maximum average intersection volume permitted for two-way stop operation be set within the range of 750 to 800 vehicles per hour. It is also recommended that, when the 85th percentile speed on the major street exceeds 40 mph, the warrants should be reduced to 70 percent of the values in Table 1.

TABLE 1

VOLUME WARRANTS FOR FOUR-WAY STOP-SIGN INSTALLATION

Peak-Period Factor	Minimum 4-Hour Average Intersection Volumes for Average Delays of		
	20 sec	30 sec	35 sec
Peak-Hour Factor = 0.75-0.80			
0.60	400	525	600
0.70	475	625	700
0.80	550	700	800
0.90	625	800	900
Peak-Hour Factor = 0.80-0.85			
0.60	425	550	625
0.70	500	650	750
0.80	575	750	850
0.90	650	850	950
Peak-Hour Factor = 0.85-0.90			
0.60	450	600	675
0.70	550	700	800
0.80	625	800	900
0.90	700	900	1,000

Notes: 1. An average delay of 30 sec per stopped vehicle is recommended for general use.
2. Intersection volumes are all-approach totals.
3. Major-minor flow ratios from 80/20 to 60/40 are included.
4. Maximum hourly volume for two-way operation is 800 vehicles per hour (4-hour average).
5. Peak-period factor equals the average hourly volume for 4 hours divided by the maximum hourly volume.

SIGNALIZED INTERSECTIONS

Traffic signals are used to assign the right-of-way alternately to vehicles or queues of vehicles passing through an intersection. For maximum efficiency, the signals should be timed so that (a) the total delay to all traffic using the intersection is minimized, (b) no individual vehicle experiences excessive delay, and (c) the average delay per vehicle is tolerable for the circumstances.

Studies of stopped-time delay at eight isolated signalized intersections, which were operated under pretimed, semi-actuated, and full-actuated control, indicated that traffic-actuated control generally resulted in less delay than pretimed control for the range of conditions observed. Apportioning of the green time was found to have a pronounced effect on delay for pretimed control. Semi-actuated control was most effective at locations where less than about 40 percent of the total traffic was consistently carried on the street equipped for detection of vehicles. Full-actuated control resulted in less delay than the two other types when the total traffic was split approximately 50/50 on the two streets or where short-time demands fluctuated on various approaches during the day.

Warrants for actuated traffic signals, which are described in the following section of this paper, were evaluated and found to provide good guidelines for selecting actuated equipment for locations where traffic volumes do not warrant pretimed signals. Delay studies at three locations that met the suggested warrants for actuated control, but not for pretimed control, showed that actuated control consistently resulted in less delay than pretimed equipment up to total volumes of about 450 vehicles per 15 min.

Studies of the effect of dial settings of actuated signal controllers on delay indicated that these settings were not extremely critical over the rather limited ranges considered to be practicable. If long loop-type detectors (40 to 80 ft long) or other suitable vehicle presence detectors that have become available since these studies were conducted are used, problems associated with detector placement, initial intervals, and vehicle intervals are virtually eliminated. Very precise controller response can be achieved by setting initial and vehicle intervals to minimum values.

An economic analysis of a representative intersection showed that the higher equipment, maintenance, and operating costs of actuated control could be easily compensated for in less than 2 years by the lower stopping, idling, and time costs that would accrue to road users from the more efficient traffic control.

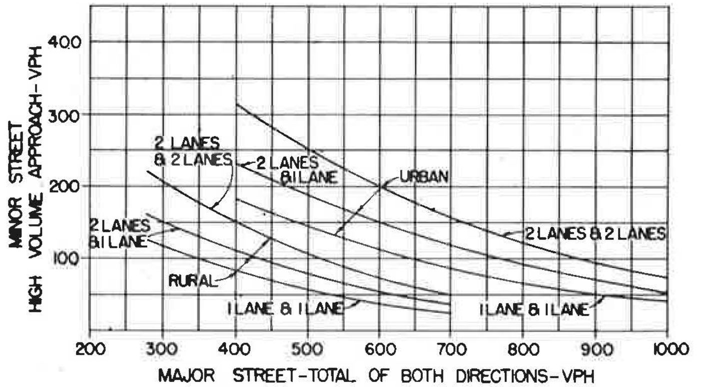


Figure 4. Warrant volumes for traffic-actuated signals, 8 high hours.

ACTUATED SIGNAL WARRANTS

A set of vehicular volume curves has been developed and used by the Texas Highway Department as a guideline in determining whether the installation of a traffic-actuated signal is justified. Volume warrants for each of (a) any 8 high hours, (b) any 4 high hours, (c) any 2 high hours, and (d) any high hour, which are shown in Figures 4 through 7 respectively, were developed as a means of analyzing the vehicular volumes, cross traffic, and peak-hour volume warrant factors for traffic-actuated signals given in the 1961 edition of the Manual on Uniform Traffic Control Devices (1). Whenever the major-street (total of both approaches) and minor-street high-approach (one direction only) volumes for each of the hours being tested rise above the applicable curve for the warrant condition, the possible installation of a traffic-actuated signal may be considered further. Although the high-volume approach on the minor street may change from hour to hour, the major-street and minor-street volumes for the same hours must be applied. The rural curves are applicable when the 85th percentile speed along the major street exceeds 40 mph or when the intersection lies within an isolated community of less than 10,000 population. The urban curves are applicable for all other conditions.

The traffic signal warrant curves for each of any 8 high hours shown in Figure 4 are based on a combined application of the volumes of pretimed traffic signal warrants 1, 2, and 6 in the Manual (1) together with a capacity curve developed by D. E. Dyas (2) for uncontrolled intersections. It may be noted that when the traffic volumes given in the

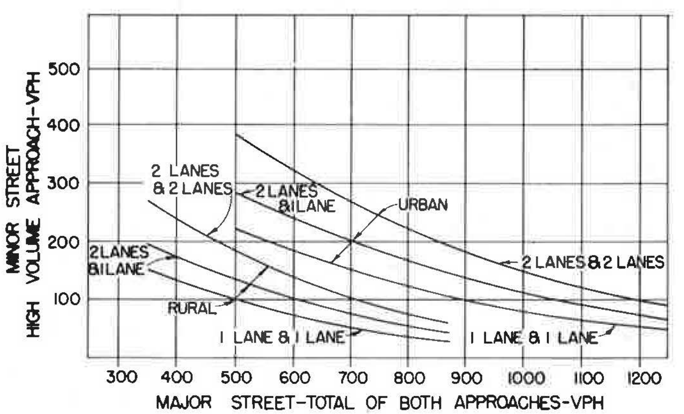


Figure 5. Warrant volumes for traffic-actuated signals, 4 high hours.

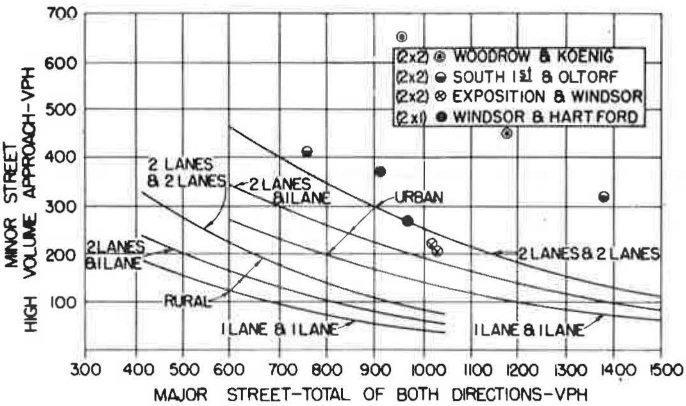


Figure 6. Warrant volumes for traffic-actuated signals, 2 high hours.

referenced warrants 1, 2, and 6 are plotted in Figure 4, the volumes fall very close to the applicable curve.

The curves shown in Figures 5, 6, and 7 represent 1.25, 1.50, and 1.75 times the volumes that comprise the curves given in Figure 4. The 1.25 and 1.50 factors are based on a traffic study conducted by the Texas Highway Department in which it was found from the vehicular volumes at 20 permanent traffic counting stations that the average hourly volume of the 8 high hours and the highest hourly volume of an average weekday were respectively 25 percent and 50 percent more than that of the 8th high hour. It was concluded that, if an intersection has traffic volumes during each of 4 high hours falling on or above a curve representing 1.25 times the volumes (Fig. 5) of those of the applicable 8 high hours curve (Fig. 4), a traffic signal could be considered further. If the traffic volumes are high during only 2 hours of a day, however, the volumes for both hours should be sufficient to fall on or above a curve having 1.50 times the volumes (Fig. 6) of the applicable 8 high hours curve. It was also concluded that if the traffic volumes are unusually high during only 1 hour of the day, these volumes should fall on or above a curve representing 1.75 times the volumes (Fig. 7) of the applicable 8 high hours curve.

The hourly traffic volumes for four study intersections having traffic-actuated signals were used in an evaluation of the curves shown in Figures 4 through 7. The results of the study show that, although none of the four intersections had sufficient traffic to meet the pretimed signal warrants set forth in the Manual, the volumes at all four intersections satisfy at least one of the urban warrants, and one intersection—South First at

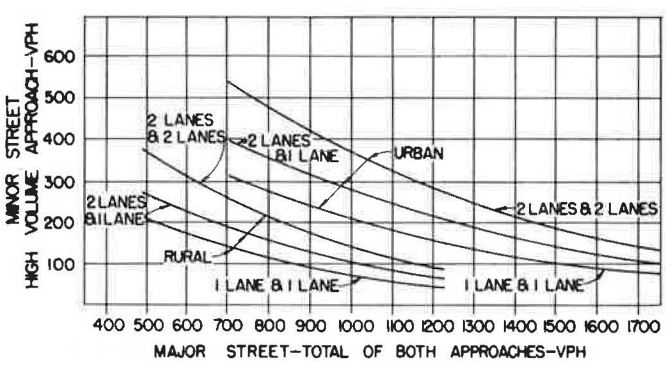


Figure 7. Warrant volumes for traffic-actuated signals, 1 high hour.

Oltorf—satisfies all four. As an example, the traffic volumes for the 2 high hours of the traffic counts for each of the four study intersections are plotted in Figure 6. As shown, three of the four intersections have sufficient volumes during 2 hours of the day to exceed the applicable warrant curve. The intersection of Exposition Boulevard and Windsor Road did not meet the warrant requirements for 4 hours. The traffic volumes at this intersection were sufficient, however, to meet the applicable warrant curve requirements for 4 high hours (Fig. 5).

SPECIAL STUDIES

The practical feasilibity of using multichannel digital recording equipment in the field for comparative delay studies was demonstrated. The recording and data analysis techniques that were developed are useful for many types of before-and-after evaluation studies. Minor modifications to the observation and analysis techniques will make it possible to use equipment similar to the D3 Recorder for studying traffic phenomena such as headways, gaps, arrival patterns, and intersection capacity.

MODERNIZED EQUIPMENT

Recent spectacular advancements in electronic instrumentation have rendered the electromechanical hardware, but not the concept, of the digital data delay recorder obsolete. Development of a new instrument system with the same basic capabilities as the D3 Recorder is recommended. It is now possible to have a portable unit the size of a small suitcase with all the features needed to conduct field traffic studies at the most complex intersections. This unit would overcome most of the limitations such as bulk, scanning rate, and complex operation associated with the D3 Recorder.

TRAFFIC SIMULATION

Computer simulation of traffic flow at intersections is potentially a powerful tool for studying intersection efficiency, but up to now very little adequate field data have been available for validating simulation models. Data collected in the traffic studies described in this report include extensive amounts of several types of information needed for verifying such models. Some of the recorded or computer information is directly applicable; other relationships can be deduced.

It is recommended that serious consideration be given to developing computer simulation models that can be used to evaluate traffic flow at isolated intersections and on street networks. Once properly verified models are available, wide ranges of traffic patterns, intersection configurations, and control techniques can be evaluated rapidly and conveniently without resorting to cut-and-try field techniques. Delay recording equipment can be used to establish quantitative information concerning realistic ranges of parameters to be evaluated by simulation.

ACKNOWLEDGMENTS

This investigation was conducted at the Center for Highway Research, The University of Texas at Austin. The authors wish to thank the sponsors, the Texas Highway Department and the U. S. Department of Transportation, Federal Highway Administration. The opinions, findings, and conclusions expressed in this publication are those of the authors and not necessarily those of the Federal Highway Administration.

REFERENCES

1. Manual on Uniform Traffic Control Devices for Streets and Highways. U. S. Bureau of Public Roads, June 1961.
2. Dyas, D. E. Traffic Engineering and the Highway Engineer. Traffic Engineering and Control. Vol. 5, No. 3, July 1963, pp. 171-175.

MEDIAN VISIBILITY IMPROVEMENTS: NEEDS, METHODS, AND TRENDS

Jason C. Yu, Virginia Polytechnic Institute and State University

In order to provide optimum traffic guidance, the median on the divided highway must remain in definite contrast with through traffic lanes to warn drivers of its presence and to delineate its perspective of geometric change at all times. Traffic accident studies, however, indicate that there is inadequate visibility of general median locations, especially during the hours of darkness. Although a variety of techniques has been developed to increase median visibility, no uniform specification has been adopted throughout the nation. It was the intent of this study to assemble and discuss various possible means of increasing median visibility through application of traffic control devices. A two-part study was undertaken: first, an intensive review of published sources was made to gain a thorough knowledge of the historical development of median visibility considerations, and, second, a national questionnaire survey was conducted to gain an awareness of current practices and any innovations for effective median delineation. As a result of intensive investigations concerning various aspects of this subject, suggestions for median visibility improvements are presented based on the current available data.

•TRAFFIC safety and accident prevention continue to assume a more prominent role in the various aspects of highway design. With the development of faster automobiles, the emphasis on safer roads is becoming paramount. The whole concept of traffic safety is based on the fact that accidents are reduced substantially if both hazardous conditions and careless actions are properly diagnosed and corrected. Because hazardous conditions are essentially constant in character, they are obviously easier to correct. Therefore, highway improvement tends to reduce highway fatalities, resulting in a safer transportation system.

Intensive investigations reveal that the foremost factor in the circumstances surrounding traffic accidents is roadway visibility. More significant is the fact that, of all the different types of accidents, vehicle head-on collisions form a prominent part (3, 5). A widely used method for reducing head-on collisions on highways is the provision of medians. Median barriers dividing highways into 2 one-way roadways eliminate head-on collisions almost entirely and often reduce other types of accidents. In general, motorists fear head-on collisions; therefore, the median, whether it is concrete or a strip of vegetation, provides a means of physical and psychological protection for the motorist against such an accident. In addition, with the use of wide medians traffic noise and headlight glare are reduced, thus resulting in less driving tension for the motorist.

Although the median provides mental and physical comfort and a feeling of security for the motorist, one of the most important functions of the median is to delineate the left side of the roadway. Therefore, to keep traffic operations safe and efficient, the physical alignment of the median must be recognized by vehicle operators at the earliest possible moment. A highway median can become more of a hazard than an aid unless it is plainly visible at all times. The satisfactory level of median visibility should, therefore, be considered as an essential safety feature of the highway facility. Just as

Sponsored by Committee on Traffic Control Devices.

the median itself is considered a safety factor in its function as a barrier between opposing traffic, increased visibility improves the function of the median as a regulator of traffic flow on the divided highway.

The nature and frequency of traffic accidents involving median encroachments indicate that better median visibility is needed, particularly during the nighttime hours. As part of this study, the result of a national survey indicates that the frequency of median accidents varies from state to state, ranging from 1 to 9 percent of the total accidents on divided highways. The accident rate is evidently deemed serious enough to warrant improved techniques in trying to curb this problem. Accordingly, effective methods of increasing median visibility must be determined and should be implemented on a national basis.

Over the past 15 to 20 years there has been extensive research in the field of highway visibility relative to greater vehicular and pedestrian safety. However, the effective measures of improving median visibility have not been rigidly established on a national basis; they have been left to individual discretion. Therefore, if there is a need to improve the visibility of some median locations, no standard method is readily available for adoption under various conditions. This indicates that specifications for median visibility under every possible situation are urgently needed and should be adopted uniformly throughout the nation.

STUDY OBJECTIVE AND SCOPE

The objective of this study was twofold. First, pertinent published sources concerning the median visibility problem were reviewed with the objective of summarizing various possible means of improving median visibility. Second, the methods currently used for median delineation and any new innovations suggested by various localities were analyzed. It is hoped that the methods collected in this report will serve as an aid in determining standard techniques that will effectively remove median visibility problems at many divided highway locations.

In order to obtain knowledge of current practices of median delineation techniques, a questionnaire was prepared and distributed to all state highway departments. The questionnaire was designed to inventory the existing and proposed utilization and application of traffic control devices for median visibility improvements. The number of questionnaires returned was 38, which represents 75 percent of a total of 51 questionnaires distributed. It should be pointed out that the results presented in terms of percentages will not consistently be a percentage of the same total number. Although 38 questionnaires were used, all were not necessarily complete; therefore, the percentage of the total used to reflect trends will be based on the number of states that responded to the particular survey question being analyzed.

MEDIAN VISIBILITY FACTORS

Before a review of various methods of increasing median visibility can proceed, it is important to identify the principal factors that exert a significant influence on the median visibility. Because of the complexities involved and the lack of sufficient data in some cases, it is not possible to include explicitly all possible factors. However, some obvious factors can be easily identified and thus will be briefly discussed.

Difficulties with median visibility may be considered to be due to four primary factors: atmospheric conditions, headlight glare, topography, and median design. Inclement weather conditions, such as rain and fog, reduce the driver's sight distance and thus significantly limit the visibility of medians. While driving at night, the glare caused by oncoming vehicle headlights, in addition to causing ocular discomfort, reduces median visibility and enhances the possibility of colliding with it. On highways with sharp curves or steep hills, the driver's attention may be distracted by the terrain; this increases the danger of hitting the median. Furthermore, various types of medians possess different degrees of visibility properties. Paved, narrow, nontraversable medians usually do not provide an effective contrast in color and texture with the traffic lane surface.

In addition, median visibility in daylight is better than that during the dark hours. This difference in visibility becomes more accentuated when special adverse factors, such as headlight glare, are involved in the night visibility of medians. Therefore, it is more important to develop and utilize delineation systems that will increase night visibility and thereby provide drivers with optimal information about the median alignment during the dark hours. In some cases, particularly at night, drivers have more difficulty in realizing that a highway is divided and not bidirectional than they have in actually seeing the median. This lack of awareness causes wrong-way driving and creates potential accidents.

MEDIAN VISIBILITY IMPROVEMENT TECHNIQUES

Many types of traffic control devices can possibly be employed to combat the multifaceted problem of median visibility. Usually several methods are used together to resolve the total problem of median visibility. The methods may be divided into groups, according to the types and functions of traffic devices used. An attempt has been made to specify the idea involved in each method as briefly as possible. Detailed specifications can be found in the source documents listed in the references, particularly in the 1971 revision of the Manual on Uniform Traffic Control Devices (6). There is no separate section written exclusively on the method of median visibility improvements in the Manual; however, it is felt that many specifications for use of traffic control devices can be assimilated to form various sections for such a purpose.

MEDIAN ILLUMINATION—METHODS AND TRENDS

A comparison of the median accident rates for the illuminated and nonilluminated sections indicates they are in the ratio of 1:2 (2). Thus, the use of highway lighting results in appreciable economic benefits due to reduced accident frequency. However, even with the substantial reduction of accidents, the erection of a continuous line of lights is still too costly. Therefore, only limited median sections of highways should be given priority for lighting.

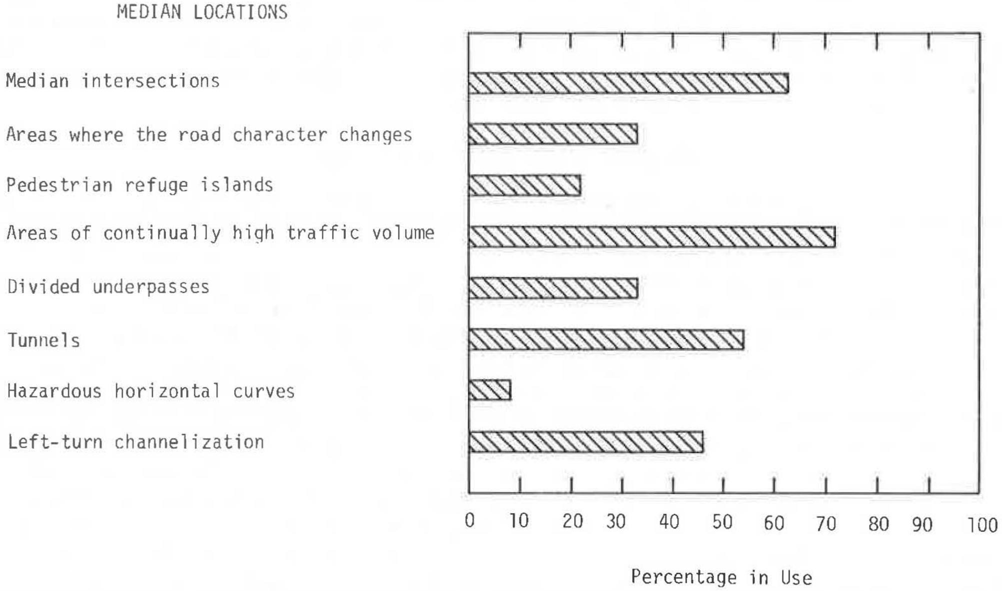


Figure 1. Hazardous median locations where lighting is used.

In the questionnaire survey a list of eight of the hazardous median locations was included, and the states were to check each location where lighting was used within their particular jurisdictions. The results are shown in Figure 1. It should be noted that the figures were compiled from the results of the 25 state highway departments that responded to this question.

Many states felt that lighting was one of the most effective methods of improving median visibility at night but that its cost prevented its widespread use, as indicated previously. The design of the lighting may fluctuate with roadway conditions; however, most states indicate that its purpose is to illuminate adequately the entire roadway in addition to the median. The method used by many states was illumination at intersections, with the light conforming to the standards of intensity and uniformity as established in the American Standard Practice for Roadway Lighting (3).

About two-thirds of the questionnaire returns suggested that under normal conditions lighting poles should not be located in the median of a divided highway with overall pavement and median widths up to 100 ft unless it is economically impractical to light the median from the outside shoulders. Where lighting poles are located within a barrier median, the total median width should be at least 8 ft and the width between barrier curbs at least 4 ft. The majority of states that responded to this question try to keep light poles off of medians in an effort to reduce accidents.

The mounting height of lighting for median visibility is recommended to be between 30 and 50 ft with 40 ft being used by a large percentage of the states. Most states also use lighting ranging between 400 and 1,000 watts, with the 700-watt luminary being predominant. The spacing between the poles is determined by the wattage of the luminary, its height on the pole, and the amount of lighting desired in that area.

APPLICATION OF DELINEATION DEVICES

Reflector Markers

Reflector markers, made of single or clustered buttons or small panels covered with reflective coatings, can be utilized to delineate the general alignment of medians. Because of the difference in the physical and optical properties of various types of reflectors, suggestions for using the different types of reflector markers for the median delineation are given in the following.

Delineators—Delineators can be used effectively to signify the median during the nighttime hours by functioning as edge markers rather than warning devices per se. They may be used on continuous sections of highway or through short stretches where there are changes in alignment. They also may be used at the approach-end of a median. The spacing is from 200 to 528 ft along relatively straight portions of the median, with closer spacing on curves (6). On the approaches to and throughout horizontal curves, the spacing should be such as to make several delineators always visible along the curve ahead of the driver. In any case, delineation should be carried through the transition on a highway with continuous delineation on either or both sides. To be effective in delineating medians, delineators should be mounted at a proper height depending on the type of delineators used. Those that separate traffic in opposite directions shall be yellow, and their spacing shall be the same on both sides of the median. Delineators should always be placed at a constant distance from the edge of the roadway.

Hazard Markers—Hazard markers may consist of single reflectors, clusters of reflectors, or small panels of uniform shape covered with a reflective coating and mounted on separate posts. Reflective elements for delineators should have a minimum dimension of approximately 3 in. In order to mark the median barriers, the delineators should be in line with or inside the innermost edge of the median and placed at a constant distance from the roadway (6). They should also be placed so as to be clearly visible to the approaching driver under normal weather conditions.

Traffic Button—A series of small reflectorized buttons may be used on the surface and near the edge of the median for effective delineation. In general use are glass and plastic buttons, molded plastic with a beaded surface, and prismatic buttons with a transparent acrylic plastic face. To define medians large buttons or bars of cast iron or concrete several inches high, with or without reflectors, lights, symbols, or messages, may also be used.

Curb Marker Light—Another method of median delineation makes use of pavement inserts and curb marker lights. A system of small, surface-mounted lights can be used to develop lineal patterns to increase the intensity of brightness and contrast at night and in conditions of inclement weather, such as rain or fog. They can be affixed directly to the median curb with an epoxy resin. The spacing must be uniform between units, except at intersections where gaps are left in the lineal patterns. Curbed sections should have closer spaced units to maintain the lineal continuity (8).

Markings

The approach end of a median in the line of traffic flow should be designed and marked to indicate its presence and also to outline the proper vehicle paths. The median end first approached by traffic should be preceded by gradually widening marks or roughened strips on the pavement designed to channel the traffic flow into natural paths of travel along the median edge. In other words, this approach-end treatment should guide the driver, without physical restraints, into the proper maneuver well in advance of the median.

Median approach pavement markings consist of a diagonal line, or lines, extending from the center or lane line to a point 1 or 2 ft to the right side of the approach-end of the median. The length of the diagonal marking should be determined by the 85th percentile traffic speed in miles per hour multiplied by the width of the offset in feet. The approach nose should be offset by a greater amount than the offset of the side of the median itself. As faced by the approaching traffic, the approach nose should be offset to the left, the right curb of the median forming a diverging taper to guide the traffic toward the right.

The 1971 MUTCD recommends that double yellow edge-marking lines be used to form median islands. This will aid visibility and also make the driver aware that the median strip is separating travel from the opposite direction. This double yellow median strip was found effective in reducing the number of head-on collisions and other types of accidents (1, 11). For the existence of a median island, a single yellow strip is used and generally placed near the edge of the median traffic lane to take advantage of the cleaning action provided by moving vehicles. If the median barrier is located in the line of traffic flow, reflectorized solid yellow paint should be placed on the curbs to channelize traffic to the right of the median.

Signs

The traffic sign is the most common device for controlling, safeguarding, or expediting traffic. For delineation purposes, signs must be used only where necessary and justified by highway conditions, such as hazards that are not self-evident or clearly visible to vehicle operators. In case of a hidden median, such signs as "Keep Right" or "Divided Highway" should have excellent visibility characteristics to provide adequate advance warning to the driver.

The message delivered by these signs, important enough during the daytime, becomes even more important at night when median visibility is limited. Therefore, the sign must be properly illuminated and reflectorized. When multiple tasks are involved, as in ordinary driving, both objects and signs can be expected to be seen at less than their legibility or visibility threshold distance. Under these circumstances, factors increasing the relative attention value of these signs are of great importance.

Most frequently the "Keep Right" sign is used to direct traffic around a median and, in particular, at the beginning of the median. Because it is of a generally standard form and widely used, the sign can probably be considered more as a symbol than a printed message. The "Keep Right" sign should be mounted a proper distance beyond the approach end of the median. The "Divided Highway" sign can also be used to give drivers advance warning of a section of the highway where the opposing flows of traffic are separated by a median. This diamond-shaped warning sign would normally be placed well in advance of the median.

Glare Screens

When the lights of an approaching automobile remain on high beam during the passing maneuver, most drivers are blinded by the dazzling light and are unable to observe clearly the median within the limits of their own headlight illumination. In various tests the effects of glare on night visibility were measured. It was shown that a large decrement in tracking accuracy occurs due to the overall effect of glare from the lights of the simulated approaching vehicles. Thus, in connection with median visibility at night, glare should be considered a potential hazard.

Because it is undesirable to reduce headlight intensity because of the accompanying loss of road illumination, it is necessary to either increase road illumination by some means without increasing glare or reduce glare without reducing road illumination. A common method used to compensate for the glare reduction that undersized medians do not provide is the use of glare screens. Several forms of glare screens are currently used.

Planting—Median plantings provide drivers with some relief from the headlight glare of opposing vehicles and provide the appearance of a third dimension of depth for roadways that seem to have only two dimensions—length and width—particularly during low-visibility conditions. Proper planting of dense shrubbery on medians can form an effective glare screen and has two additional beneficial effects: the shrubs provide some degree of cushioning for vehicles tending to cross into the opposing traffic lane, and the shrubbery adds to the natural appearance of the scenery (10).

Guardrails—Although the main purpose of guardrails is to prevent intrusion into the opposing lane, they do afford some screening of glare that results in increased median visibility (4, 9). The guardrails seem most effective on separated highways where the opposing lanes are at different elevations. Because the general character of the headlight is to direct light downwards, the glare from the higher road surface is produced with both the high and low beams of a vehicle's headlights. In this situation, a guardrail serves a double purpose when placed on the upper roadway. Besides preventing the intrusion, it shields the eyes of the driver from harmful glare when placed at an appropriate height, depending on the difference in elevation of the two roadways. To be fully effective, guardrails must be highly visible, well-maintained, painted assemblies with reflector buttons or reflectorized materials. A study (2) was conducted to determine the effect of median guardrail on accident rates and severities. The result indicated that median rails can increase the accident rate but the overall severity rates were lowered.

Wire Mesh—The most satisfactory glare screen has been found to be a line of expanded metal mesh placed on the median, parallel to the centerline. Its effectiveness as a screen results from its manufacturing process. The strands of the screen have twists that block any light traveling tangent to the face of screen. The formation of the strands allows the fence to become transparent at angles greater than 20 deg (when glare is relatively unobjectionable). During the day the fence does not block the general view of travelers and allows police surveillance of the opposing traffic lane.

The height of the screen is determined by the use of several relevant factors. The width of the median, the height of the headlight, and the height of the driver's eyes all determine the location of the upper and lower edges of the screen. The obstruction of the glare of high-beam headlights is mandatory. The screen must be low enough to prevent the lights of the smallest sports car from blinding the drivers in the opposing traffic. Conversely, the screen must be high enough so as not to allow the lights of large trucks and buses to shine into the eyes of truck and bus drivers in the opposing lane. The metal screen is not practical along sharp curves or narrow medians because of damage from the overhanging parts of trucks (7).

TRENDS OF DELINEATION TECHNIQUES

All of the state highway departments were asked to list the types of traffic control devices being used for improving median visibility. The method used, of course, is dependent on the conditions under which the median visibility problem exists. However, some of the common techniques used today are shown in Figure 2.

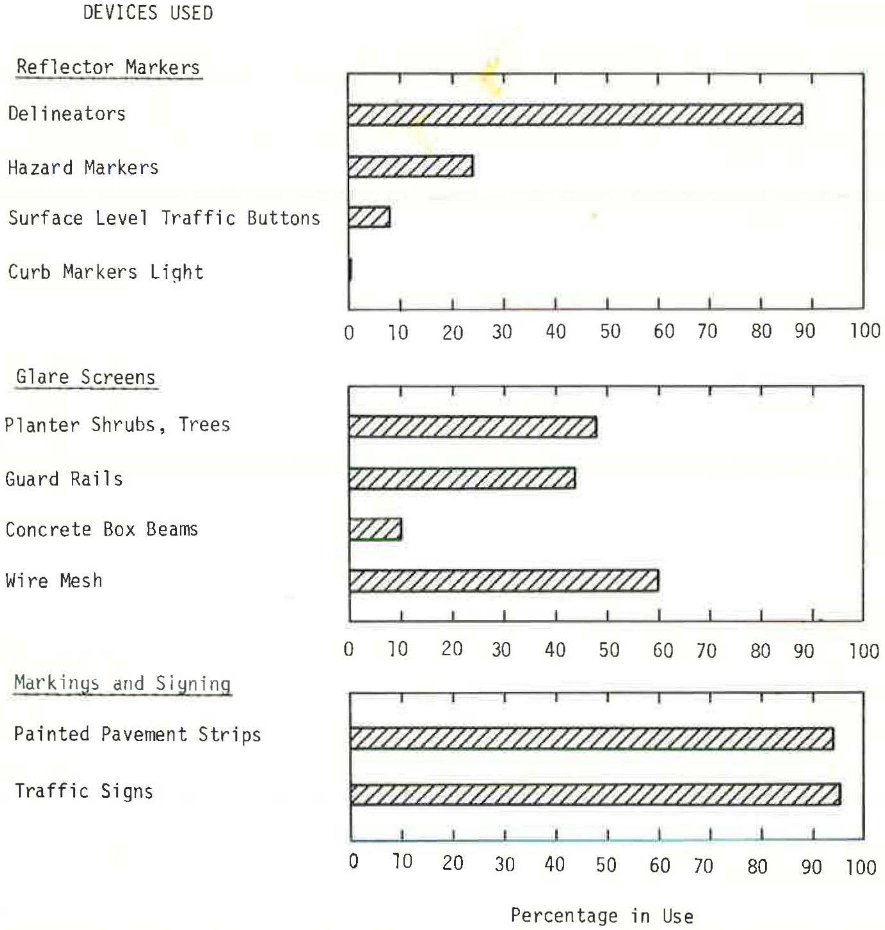


Figure 2. Median delineation methods used in states.

A total of 24 states answered this question. The census showed that pavement markings have generally been used on narrow medians in most of the states. Delineators and hazard markers are used on narrow medians only in an attempt to achieve recognition of the presence of a raised median, particularly at night. Glare screens or guardrails are most common and are used only on narrow medians where traffic volume is high. Planted shrubs and trees are used on curbs but are generally used to block headlights rather than to serve as delineation. Signs in medians are restricted to protected locations, and most states use breakaway sign posts for safety reasons. The percentage of accidents involving collision with median signs was found to be relatively small, varying from 1 to 6 percent from state to state. It is also indicated that most states are presently interested in using the wire mesh glare screen. Fourteen states are currently conducting experiments to test the cost-effectiveness of this technique for the reduction of glare.

Another question was asked concerning the dominant factors in choosing which type of median delineation method to use. The results indicate the following priorities of importance:

1. Economical in installation and maintenance;
2. Compliance with past practice;
3. Aesthetically pleasing;

4. Eliminate, if possible, fixed objects or sight distance obstructions in the median; and
5. Form a continuous marking of the median edge as well as lane control.

VERTICAL DIMENSION TECHNIQUES

With the high volume of traffic that the driver faces on divided highways and the lack of objects in the median to create the third dimension of height, the driver in the median lane tends to watch traffic to locate his position. Thus, with no easy means of determining his lateral position or speed by reference to an object at eye level, the driver has to divert his attention from traffic to view the curb, guardrail, grass, or edge line that delineates the median. This constant visual activity required by the driver to orient his lateral position increases the potential for median encroachments and possible encroachment into the opposing traffic flow.

Many devices, such as curbs, guardrails, fences, reflectorized delineators, raised markers, and painted lines, as well as the general appearance of the median, provide the delineation of the left extremity of the roadway during both daytime and nighttime hours. However, if these devices do not provide the dimension of height as well as of length and depth, they cannot be used efficiently or effectively by the driver to make quick judgments concerning his lateral position. During periods of limited visibility created by rain, snow, fog, or curves in roadway alignment, these delineators become even less effective.

The survey asked a question in relation to the use of vertical dimension techniques. In response, the highway departments (26 returns) listed the techniques now in use to separate the two opposing flows of traffic. Six techniques were listed, and a space for other methods was included. The results of this aspect of the survey are shown in Figure 3.

The three techniques that are most commonly used in rural areas are the split-level roadway design, evergreens or trees, and woodland areas. In essence, the highway departments are constructing two separate pavements with the area between them left in its natural state.

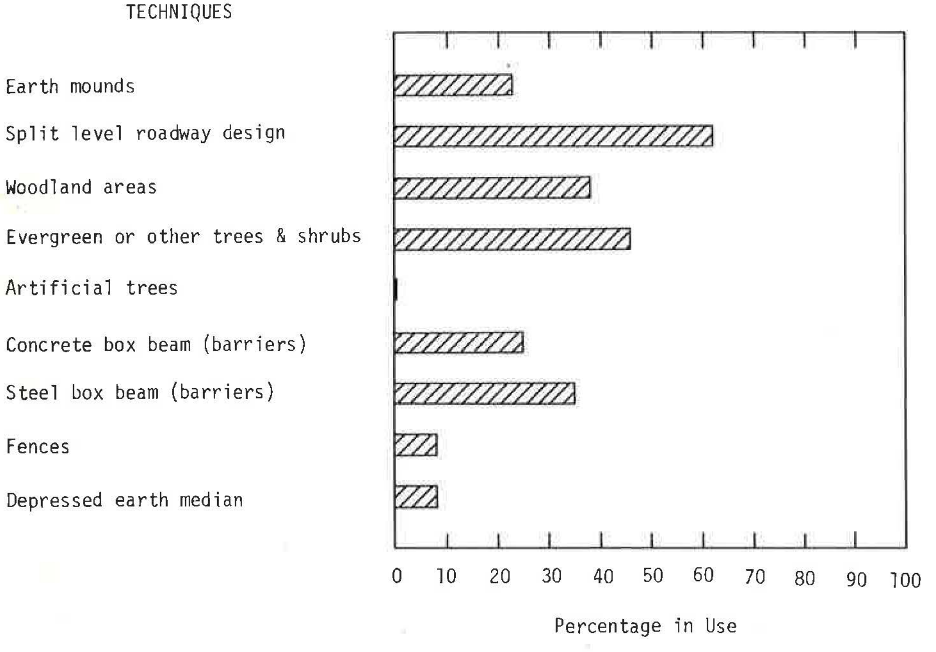


Figure 3. Vertical dimension techniques.

The highway departments were also asked for the dominant factors considered in choosing the vertical dimension techniques used in their states. The following list is a summary of these factors and the order of their relative importance:

1. Avoidance of headlight glare from one roadway to the other;
2. Economy in construction and maintenance;
3. Roadway geometrics and median width;
4. Urban or rural areas;
5. Traffic volume;
6. Preservation of natural topography and growth; and
7. Aesthetic appeal.

The dominant factors in rural areas include cost and the topography of the land. Urban areas also consider these factors and in addition also note the design of the highway and the physical limits of the area in which the highway is located.

If the median is going to fulfill its prime objective of delineating the left extremity of the roadway, the third dimension of height must be incorporated into the design of today's highways. The use of physical features to create this dimension has proved effective in helping the driver remain alert to driving conditions. There is still a great need for research into this field to create effective third-dimensional objects that will not become potential hazards to the motorist.

CONCLUSIONS

The highway median, separating opposing traffic lanes of a roadway, has been found to be an effective measure when used with other control means in assisting traffic to move in a safe and efficient manner. The unexpected appearance of a median can startle a driver and possibly cause an accident. Therefore, the presence of the median should be signaled far enough in advance so as not to require immediate action.

Several possible methods for increasing the visibility of highway medians have been reviewed and summarized. It is quite evident that no single method will always give the best results under all circumstances. Although the important test of any of these methods is the effectiveness of increasing median visibility, the problem of selecting an appropriate method under a given condition is partly an economic one. The resulting median visibility achieved by a method must be weighed against the cost of the method.

To be incorporated into both existing facilities and the plans of the future facilities, effective methods that will enhance the median visibility must be found and utilized correctly in order to avoid all possible traffic accidents due to inadequate median visibility on divided highways. This report has attempted to bring up to date the developments in the area of median visibility. Suggestions based on the current data have been made concerning several aspects of this topic. It is imperative that, as further research is compiled, continuing evaluations be made on this important highway element.

REFERENCES

1. Evaluation of Minor Improvements: Delineation and Guard Rail (Parts 3 and 4). California Division of Highways, 1967.
2. Expressway Median Barrier Rails. Connecticut Department of Transportation, 1969.
3. American Standard Practice for Roadway Lighting. Illuminating Engineering Society, 1963.
4. Evaluation of Three Installations of "Blocked-Out" Median Guardrail With Glare Screen. Michigan Department of State Highways, 1969.
5. Accident Facts, 1965. National Safety Council, Chicago, 1965.
6. 1970 Revision, Manual on Uniform Traffic Control Devices for Streets and Highway. U. S. Department of Transportation, Bureau of Public Roads, 1971 (Draft).
7. Coleman, R. R., and Sacks, W. L. An Investigation of the Use of Expanded Metal Mesh as an Anti-Glare Screen. Highway Research Record 179, 1967, pp. 68-73.
8. Deakin, O. A. Planting for Screening Headlight Glare and Traffic Guidance. Highway Research Record 53, 1964, pp. 17-25.

9. Hofer, R., Jr. Glare Screen for Divided Highways. HRB Bull. 336, 1962, pp. 95-101.
10. Hutchinson, J. W., and Lacio, J. H. An Experiment With Evergreen Trees in Expressway Medians to Improve Roadway Delineation. Highway Research Record 105, 1966, pp. 85-98.
11. Johnson, R. T. Median Striping Study. California State Highway Division, 1964.

DRIVER INFORMATION NEEDS

T. M. Allen, Michigan State University;
H. Lunenfeld, AIL, Division of Cutler-Hammer, Inc.; and
G. J. Alexander, U.S. Department of Transportation, Federal Highway Administration

The driving task was analyzed to determine the nature and interrelationship of the subtasks the driver performs and the information needed to perform them safely and efficiently. Data were developed using a modified information-decision-action task analysis method applied to several long driving trips. The task analysis provided the basis for categorizing the various component driving subtasks, identifying information needs associated with the subtasks and their present methods of satisfaction, and providing a structure to the driving task. Driving subtasks were categorized in accordance with information-decision-action complexity and ordered along a continuum. The subtasks were found to fall along a hierarchical scale. Vehicle control subtasks such as steering and speed control were ordered at the lowest level and identified as micro-performance (control). At an intermediate level, subtasks associated with response to road and traffic situations were identified as situational performance (guidance). The highest level subtasks, encompassing trip planning and preparation and route finding, were identified as macro-performance (navigation). Performance of subtasks at the high level of the hierarchy involves component performance at a lower level. Drivers search the environment for information needed to perform the various subtasks and shift attention from one information source to another by a process of load-shedding. When load-shedding is required due to the demands of the driving situation encountered, information associated with subtasks relative to the subjective needs of the driver is attended to, and other information sources are shed.

•THE importance of providing the driver with information needed to perform the driving task has been pointed out by Cumming (1), who states, "The road complex must provide for the operator a comprehensive display of information both in the formal sense of signs, signals, guidelines and edgeposts, and in the informal sense of clear visibility in all relevant directions." Design and placement of these formal displays must be compatible with the prevailing vehicle speeds, traffic pattern, and visual and response characteristics of the human operator. Moreover, they must be as free as possible of irrelevant, distracting material and should be able to function in spite of competing demands for attention.

The project statement of NCHRP Project 3-12 (2), which led to the research reported here, stated, "...with the ever-increasing demands of the driving task, there is urgent need for improvement in the understanding of the driver's needs for information and the means of communicating it to him."

NCHRP Project 3-12 was performed by AIL, a Division of Cutler-Hammer, and was designed to elicit answers to the following questions:

1. What information is needed by the highway users for their safe, convenient, efficient, and comfortable performance of the driving task?

Sponsored by Committee on Motorist Information Systems and presented at the 50th Annual Meeting.

2. What are the principal factors and interactions underlying the reception and use of this information?

3. To what extent can visual communication be used successfully?

The final report on this project was submitted to NCHRP (3).

The present paper reports one aspect of this research—the determination of the information needed by drivers and the principal factors and interactions underlying the reception and use of needed information.

THE DRIVING TASK

The purpose of information is to reduce uncertainty. As long as uncertainty exists, alternate possible decisions cannot be fully evaluated. In order to make rational decisions, the driver must reduce his uncertainty. The need to reduce uncertainty is generated from an information need that requires information for its satisfaction.

AIL thus started its investigation of the drivers' information needs by determining what the driver does. This analysis of the driving task gave an indication of the decisions that have to be made by the driver and, therefore, of the information required to make the decision.

An analytical method was required to determine the information needed by drivers and to provide a framework for conceptualizing the form and timing of information presentation so that it can be used most effectively. The "man-machine systems task analysis," used extensively in the analysis and design of military subsystems, was applied to this problem.

Driving Task Analysis Procedure

Task-analysis methods have had only limited application in driver-related research. A method similar to task analysis is the use of an "events recorder" by Greenshields (4), which has subsequently been used by Platt (5) and others for such uses as evaluating the effects of fatigue on driving performance. Other researchers such as Algea (6) and Biggs (7) have analyzed portions of the driving task, with particular attention to information inputs. Until recently, the only reported attempt to analyze the driving task as a whole has been that of Miller (8). However, the purpose of Miller's paper was to illustrate the technique of task analysis, using driving as an example. As such, it had neither the breadth nor attention to information inputs to satisfy the needs of this study.

Task analysis has been defined by Seale (9) as "... that portion of the total system analysis effort which defines systematically and in as much detail as possible at any given time, the stimulus inputs to the operator, the response output of the operator, and the operational environment in which he works."

On the basis of pilot runs using several task analysis procedures, the format that best served the objectives of this study was a modification of a task analysis method developed by one of the authors (10) in a military information needs and control actions study. This task analysis collected data (a) for each situation encountered in transit; (b) for each piece of information displayed; (c) on driver observations and expectancies relative to information needed; (d) reflecting road and traffic conditions; (e) on driver perception; (f) on driver cognition (evaluations, predictions, and decisions); (g) showing driver control responses; and (h) on feedback information.

This format was used to analyze verbal observations obtained from one long trip (from New York to Michigan) and several short trips on urban freeways. The data collected resulted in about 1,000 feet of tape-recorded task analysis data. Data from the tapes were analyzed and categorized into the subtasks representative of the behavior described on the tape and transposed into the form shown in Table 1. Further analysis of the data yielded the description of the driving task discussed in the following and information needs associated with the various driving subtasks.

Description of Driving Task

The driver performs a number of interrelated subtasks, some of them simultaneously. In developing an overall concept of the driving task, an analysis was

TABLE 1

VEHICLE CONTROL TASK ANALYSIS

Item		Speed (mph)	Road Conditions and Traffic Dynamics	Driver Dynamics			Vehicle Response	Feedback[b]	Distractions	Remarks	
No.	Elapsed Mileage			Perception	Cognition[a]	Response				Suggested Information Aids	Comments
1	—	—	—	—	—	—	—	—	—	—	—
1A	000.0	00	Parking lot	Rented car	EV—New car PR—All equipment operative DE—Since car rented must familiarize self with equipment	Enter car Adjust seat Adjust seat belt Adjust mirror Determine location of equipment	None	V/T	—	Gages and displays in car to show all systems operative	Predriving task
				All equipment adjusted	EV—All equipment adjusted PR—Will perform satisfactorily DE—Start car	Place key in ignition Turn key	Car starts	V/T/A	—	See above	First driver task under all situations Time: Morning Weather: Clear Driver: Alert
1B	000.0	20	Road: Single access lane feeding into "main" road Traffic: Access lane clear of traffic	Clear access lane leading into desired road	EV—Access lane clear PR—Will lead to desired road DE—Turn onto access lane	Put car into gear Turn wheel to desired direction Depress gas pedal lightly	Car moves onto access lane at 20 mph	V/T/K	Pedestrians	Directional information showing where access leads to	
	000.3	10	Road: Juncture of access lane and "main" road Traffic: Access lane clear of traffic	Yield sign	EV—Other cars have right-of-way PR—Will encounter traffic of "main" road but merge is such that entrance onto main road can be accomplished without full stop	Take foot off gas pedal and let car coast to merge point	Car slows	V/T/K	—	Description of traffic conditions ahead Directional information showing where "main" road leads Number and name of main road	Anticipatory task requiring vigilance on driver's part Could be danger point under poor ambient conditions

[a]EV = evaluation, PR = prediction, DE = decision.

[b]V = visual, T = tactile, A = auditory, K = kinesthetic.

performed of interrelationships among the subtasks. It was determined that these subtasks could be ordered into a hierarchy that describes the organizational content of the driving task. The hierarchy itself is ordered according to time scale and level of cognitive activity. The subtasks differ in the time scale relevant to their performance from fractions of a second for steering to minutes or hours for trip route finding. They also differ in the level of cognition (mental) required by the driver. The cognitive activity required for steering is nonverbal, highly overlearned by experience, and might be performed by an animal capable of operating the controls. The task of route finding, on the other hand, requires thinking in terms of abstract symbols, language, and maps.

Steering and speed control, which are continuous throughout the driving task, are involved in the performance of all higher level subtasks responding to road and traffic situations.

Figure 1 is a schematic representation of the hierarchical description of the driving task and ties together the concepts of levels of performance and primacy (horizontal axes) that will be discussed next. Figure 1 also indicates how classes of information needs (vertical axes) interact.

Levels of Performance—The vehicle-control subtasks low in the hierarchy are those observed in looking at the fine details of driving. They are referred to as microperformance, or control. Macro-performance, or navigation, refers to the large behavioral subtasks at the high end of the hierarchy. The remaining subtasks in between consist mainly of responding to roadway and traffic situations and are referred to as situational performance or guidance. Thus, this characterization in terms of behavioral subtask levels is referred to as levels of performance.

A partial listing of subtasks, in order from low to high on the hierarchy, includes steering, speed control, responses to road situations, carrying out maneuvers, bringing about a change of route, route finding, and trip planning. Analysis of the hierarchical ordering reveals that performance of a subtask at any level in the hierarchical scale affects each subtask lower in the hierarchy. For example, steering and speed control are involved in the performance of all tasks higher in the hierarchy, whereas executing a trip plan involves all subtasks lower in the hierarchy.

Attention—An important characteristic of the driving task description is that it describes which subtask the driver will and/or should attend to as a function of the demands of the situation. When performance demands are minimal for subtasks low on

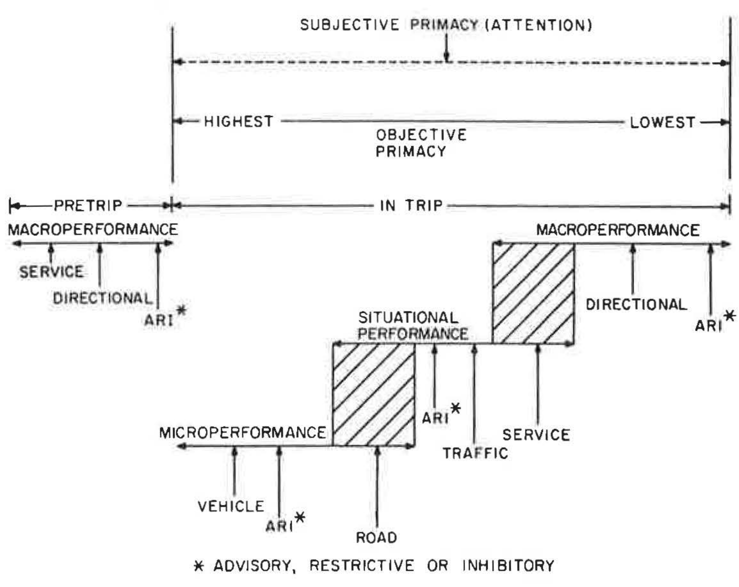


Figure 1. Description of driving task.

the hierarchy, these are performed with little conscious attention. In these cases, the driver's attention may be directed toward subtasks higher in the hierarchy. However, when high-performance demands occur for a subtask at any level, the driver should not attend to subtasks higher in the hierarchy but still must attend to those lower in the hierarchy. For example, on a straight road in free-flowing traffic, a driver can attend to route finding. However, an event on the road such as a vehicle cutting him off causes his attention to shift from route finding. He cannot, however, ignore the control subtasks lower on the hierarchy—each of these must be performed in response to the traffic situation. Thus, the hierarchy describes the "load-shedding" behavior of the driver—when the driver becomes overloaded by a subtask at one level of performance, he sheds all tasks higher but not those lower.

Primacy—The description permits establishing a priority or primacy of subtasks and their associated information needs; information needs lower in the hierarchy have priority over needs higher in the hierarchy. This priority follows from the "load-shedding" of subtasks described above. For example, if a hazardous traffic situation demands the driver's attention, he may fail to receive the message of guide signs. If the driver has inadequate information to perform a necessary control or guidance subtask that requires his attention, information relevant to subtasks higher in the hierarchy should not be presented to him at this time. Lower priority information, if presented vividly, could distract him from obtaining the higher priority information he needs.

Expectancy—An important factor in describing the driving task is expectancy. The driver has an expectation of the vehicle response to his steering movements, learned by experience. Similarly, he has expectations concerning the curvature of the roadway, behavior of other vehicles, and the signs he will find to direct him on his route. The expectancies, which apply to very short distances ahead at the low end of the hierarchy and long distances ahead for subtasks on the higher end of the hierarchy, play an important part in the integration of the driving task. When such expectancies are not fulfilled for one of the subtasks, performance of that subtask, and perhaps subtasks lower in the hierarchy, may be seriously disrupted and result in hazardous driving.

DRIVER INFORMATION NEEDS

After developing a structure for the driving task and its constituent subtasks, the information needs associated with these subtasks were organized in accordance with the hierarchical conceptualization. Although it was not possible to provide an inventory of driver information needs that took into account all possible trips and situations, it was possible to derive a list of information needs representative of the types of needs associated with typical subtasks. These needs, discussed in the context of the relevant major levels of performance categories of subtasks, are discussed next.

Control (Micro-Performance) Information Needs

There are essentially two major control subtasks at the micro level—steering control and speed control, with elements of each involved in all major subtasks in driving at the micro and situational level.

Steering Control—The major information needs associated with steering control involve vehicle response characteristics and vehicle location information and all changes thereof.

The following is a simplified discussion of the steering control subtask indicating how the information required for the task is used by the driver.

In the simplified case of a flat, straight road, the driver only requires lateral position information with respect to the road to apply minute steering corrections and thus maintain a "steady state" (11). The time frame for such vehicle control activities is on the order of ½ second. For the horizontal curve, larger control actions are required to maintain lateral position commensurate with the geometrics of the curve, and several seconds may be required to track the curve.

The steering subtask requires the driver to maintain spatial orientation with respect to the roadway immediately ahead of him. Because this task is lowest on the hierarchy, this information is of highest priority. Thus, one of the most basic information needs is location and relative lateral movement with respect to the roadway. Gordon (12) and Michaels and Cozan (13) indicate that perception of road edges and lane divisions satisfies these needs. Feedback of the vehicle's response to steering wheel movements is also necessary.

The driver-vehicle-road system has been viewed as analogous to a closed-loop servo-mechanical system by Algea (6), Rashevsky (11), and Biggs (7). The driver receives visual feedback of changes in position and orientation with respect to the road. He also receives "seat of the pants" feedback (kinesthetic and tactual), associated with centrifugal acceleration of the vehicle and lateral slope of the highway. If centrifugal force is sufficiently high to be near the frictional resistance of the tires, tire noise may inform the driver of an impending skid.

Information on changes in vehicle response is important when high demands are placed on the steering task. The driver learns by experience the responses of the vehicle to his steering wheel movements (the handling characteristics of the vehicle) and integrates these with his perception of position, lateral motion, and orientation. The steering required in most driving situations is heavily overlearned and can be carried out without conscious attention. However, emergency situations might require steering performance beyond the limits of the driver's experience. Speed (velocity) and changes in speed (acceleration and deceleration) are the main information needs obtained by visual perception of the environment, aided by the speedometer. Tactile and kinesthetic perception of accelerator and brake pedal and vehicle response is also an information reception channel.

Speed Control—To maintain a steady speed on a flat road, the driver's task is comparable to steering on a straight road, requiring only minor control actions. Changes in vertical alignment require larger changes in accelerator and/or brake use (14). Barch (15) found that drivers could make accurate speed judgments without a speedometer. Consistent overestimates or underestimates were associated with the speed to which the driver was adapted. Barch failed to find the "velocitation" (underestimates of low speeds associated with long periods of driving at high speeds) which is commonly believed to exist. Denton (16) attempted to establish a subjective scale of speed. He found consistent errors in speed judgment associated with adaptation, environment, day versus night, and physical state of the driver.

Speed control has also been characterized by a servo-mechanical model, with the data of Todosiev et al. (17) generally congruent with such a model.

Speed control affects steering, even in simple cases, as shown by Wohl (18). The speed of the vehicle on a given horizontal curve affects steering greatly, and braking to the point of skidding makes steering control almost impossible. Vehicle control requires integration of the two tasks and anticipation of imminent vehicle control needs. In addition to maintaining the desired speed, the driver must observe changing conditions and respond so that he will arrive at every point in the traffic situation at a speed at which he can control his vehicle safely on the desired path.

Speed control requires attention further down the road than does steering. In addition to visibility of the roadway sufficient to bring the vehicle there at a "safe" speed, any conditions of the roadway not consistent with driver expectancy represent special information needs. Warning signs to some extent currently provide for these needs.

Guidance (Situational Performance) Information Needs

Whereas the control subtasks are very limited, the guidance subtasks at the situational level are as varied as the number and types of road and traffic situations encountered. Therefore, only a few examples of situational performance subtasks are presented. However, as a general rule it can be stated that the information needs at this level involve information relative to all aspects of the highway system, such as other cars, road geometrics, obstacles, and weather conditions.

Vehicle guidance may be characterized as the object of maintaining the most efficient and safe course, relative to static and dynamic factors in the environment that are generally beyond the driver's control. Performance at this level is a function of the driver's perception of a situation and his ability to respond in an appropriate manner. Therefore, the driver must have a store of a priori knowledge on which to base his control actions as well as an understanding of what the situation demands. Some examples of situational performance subtasks are as follows:

1. Car following—In car following, the driver is constantly modifying his car's speed to maintain a safe gap between his car and the vehicle he is following. Thus, in this situation he is time-sharing tracking with speed control activity. He has to know, as a minimum, lead car speed, changes in its speed, how fast he is traveling, and the relative distance between his vehicle and the lead vehicle.
2. Overtaking and passing—A second guidance subtask that commonly occurs is passing, which involves, in addition to speed control, modifications in the basic tracking activity. In passing, the driver is required to know control information, such as how fast the lead car is traveling and the acceptable gap. He must, in terms of control actions, know how to maneuver his vehicle so as to use the adjacent lane gap most safely.
3. Other situational subtasks—Among the guidance subtasks that may occur are avoidance of pedestrians, response to traffic signals, advisory signs, and other formal information carriers such as stop signals at railroad crossings.

In all cases, the important point, in terms of information needs, is that the driver must receive information so that (a) he is aware of the occurrence of a situation, and (b) he knows what the situation is. Furthermore, he must possess the skills and a priori knowledge that will enable him to make the appropriate steering and speed control responses. He should also have information (feedback) that will indicate the adequacy of his response.

The driver is constantly time-sharing subtasks at this and lower levels. Two or more guidance subtasks may occur simultaneously or in close time proximity. Moreover, control actions may not be compatible, again pointing to the importance of experience, skill, and a priori knowledge throughout the driving task.

At the situational level, the driver must scan the environment and obtain information from many sources (19) to maintain an appreciation of a dynamic situation. He must also rely on judgment, prediction, and estimation, as well as feedback, to maintain what Schlesinger and Safren (20) characterize as an "area of safe travel" relative to his car and the elements of the highway system. Although elements of cognitive behavior at the guidance level are similar to those at the control level (6), a higher level of decision-making is required for most guidance tasks.

Information needed by the driver at the guidance level is that which enables him to maintain a complete appreciation of all events that could possibly affect his safe travel. Thus, he needs information on the relationship of his vehicle to the road, other vehicles, and the environment.

Studies on situational information requirements have included research on the ability of drivers to detect the speed and gap of other cars. Olson et al. (21) found that drivers were accurate in determining whether the distance between their car and a lead car was increasing or decreasing but that they tended to underestimate the relative speed differential between their car and the one in front of it. Braunstein and Laughery (22) found that drivers responded to the occurrence of acceleration and deceleration rather than the magnitude.

Several studies [Hoppe and Lauer (23) and Stalder and Lauer (24)] concerned the perception of motion between vehicles under reduced and night visibility conditions. These showed that better visibility makes it easier to perceive whether a car is coming toward you or going away from you and also that the greater the speed the more difficult it is to perceive speed, all other things being equal.

Several studies were directed toward gap and following distance. Wright and Sleight (25) discussed the "rule of ten," calling for one car length spacing for each 10 mph of speed, and characterized it as being unrealistic. A study by Lerner et al. (26) attempted to determine how following distance on the highway was affected by day versus night, trip duration, traffic, and speed. They found that the only factor of the four tested that affected following distance was speed, with greater following distances found at higher speeds. Several investigators attempted to provide displays to give supplemental gap information. Bierley (27) tested two types of vehicle spacing visual displays. One provided the actual distance between the driver and a lead car, and the other provided the algebraic sum of the gap and the relative vehicle speed. He found that the latter display increased spacing stability. Fenton (28) proposed a tactile display for gap information that used the same principle that Bierley's algebraic summing display used. Both of these displays offer significant improvement over the present unaided means of determining gaps.

Another group of studies was directed toward the ability of drivers to make judgments in passing situations. Bjorkman (29) reported, in an experiment design to determine how accurately a driver is able to estimate where he will meet an oncoming car, that subjects made errors toward the midpoint between the two cars rather than the actual meeting point.

In a study by Crawford (30) designed to determine how well a driver can decide whether to pass a lead car, he found that in over 8 percent of the cases the drivers were wrong. Jones and Heimstra (31) performed a study to determine how accurately drivers could estimate the "clearance time" required to pass another car (by "clearance time" they were referring to the last possible moment that drivers could make a passing maneuver). They concluded that drivers could not make this judgment accurately.

A study by Brown (32) indicated that a car radio seemed to have a beneficial rather than a detrimental effect on driving in both "light" and "heavy" traffic. A study by Hulbert (33) attempted to determine whether driver galvanic skin response (GSR) could be used to record traffic events; his results indicated that they could not be used. A similar finding was made by Taylor (34), who attempted to correlate drivers' GSR's and accident rate, with negative results.

Finally several studies on intervehicle communication were performed. Gibbs et al. (35) compared the European illuminated mobile arm turn indicator with the U.S. flashing lights, with results showing the U.S. system to be superior. Shor (36) studied nonverbal expectations as a means of intervehicular communications and concluded that confusion in driving in traffic results from misinterpretation of other drivers' intentions due to different driving patterns in different locations.

It is seen that considerable study of the perceptual and cognitive factors associated with the various situational subtasks and their integration is required. Similarly, the information needs associated with these subtasks must also be more precisely determined.

Navigation (Macro-performance) Information Needs

To fully describe the driving task, the third level of driving performance—the navigational level—must be considered. This level takes into account the way in which the driver plans a trip and executes his trip plan in transit. Thus, the macro-performance level consists of two phases: trip preparation and planning, which is usually a pretrip activity, and direction finding, which occurs while in transit.

Trip Preparation and Planning—Drivers use various means to formulate trip plans depending on experience, pretrip sources, and nature of trip. The means can be as formal as having the trip planned by a touring service or as simple as using a familiar route. It may consist of a driver reading existing maps and formulating the trip on his own receiving verbal instructions, or having a conceptualization, however vague, of where his destination is in relation to known routes and past experience, with the driver hoping for directional signs that will lead him to his destination. However minimal the preparation, it is unlikely that a driver will attempt to get to some destination completely unprepared.

The results of a direction-finding analysis have shown the importance of good trip preparation. The better prepared the driver is, the easier will be his direction-finding task, regardless of how poor the in-trip directional information is.

Direction Finding—During the direction-finding phase, the driver on the road must find his destination in the highway system in accordance with his trip plan and the directional information received in transit. He must thus share navigational subtasks with subtasks at the other driving levels. The navigational subtasks are further complicated because the information needed at this level, although essentially directional, may include consideration of such things as availability of services, availability of alternate routes, etc. Needs of the driver and/or his passengers that may arise in transit also affect the driver's information needs at this level.

Conversely, micro-performance and situational performance factors can also affect the macro-performance level. For example, a vehicle malfunction can lead to the

TABLE 2

DIRECTIONAL MACRO-PERFORMANCE INFORMATION NEEDS

Item	Information Need	Definition	Present Means of Reception and Transmission
1	Directions to intermediate destination[a] (type need)	Information telling driver how to find his way to an intermediate destination (stop-over, rest area, city along the way, interchange, etc.)	Visual—perception of signs (LONG ISLAND EXPRESSWAY NEXT EXIT) A priori—pretrip mapping, oral instructions In trip—determined by asking someone in transit
2	Directions to final destination (type need)	Information telling driver how to find his way to final destination (end of trip)	Visual—perception of signs (NEW YORK CITY STRAIGHT AHEAD) A priori—determined by maps, oral instructions, etc. In transit—determined by asking someone in transit
3	Distance to intermediate destination	Indication to driver of how far (in road miles) he must travel to arrive at his intermediate destination	Visual—perception of signing (NEW YORK 90 MILES) A priori—knowledge of distance from map
4	Alternate route; overall (type need)	Indication of different routes available to arrive at destination	Visual—perception of signs (NEW YORK VIA PARKWAY OR EXPRESSWAY) A priori—determined by prior mapping In transit—determined by asking someone in transit
5	Alternate route; segment (type need)	Indication of alternate routes available in the event of tie-up	Visual—perception of signs (ALTERNATE ROUTE TO BROOKLYN NEXT EXIT) A priori—prior knowledge of alternate route Auditory—commercial radio
6	Designation; road name/number (specific need)	Indication of road name and/or number	Visual—perception of signs (US 1) A priori—pretrip determination
7	Designation; interchange (specific need)	Indication of interchange name and/or number	Visual—perception of signs (EXIT 41) A priori—pretrip determination from maps, etc.
8	Designation; entrance (specific need)	Indication of entrance name and/or number	Visual—perception of signs (ENTRANCE TO INTERSTATE 95 NORTHBOUND) A priori—pretrip determination from maps, etc.
9	Designation; exit (specific need)	Indication of exit name and/or number	Visual—perception of signs (EXIT 17—NEW YORK) A priori—prior knowledge from maps, etc.
10	Designation; turn-off (specific need)	Indication of turn-off name and/or number (point other than an exit, entrance, or interchange)	Visual—perception of signs (ENTRANCE TO HOLIDAY INN PARKING LOT) A priori—determined from maps, etc.
11	Elapsed mileage (type need)	Indication of distance traveled (from some reference point)	Visual—perception of odometer Visual—perception of mileposts
12	Distance to final destination (type need)	Indication of miles to go to destination	Visual—perception of signs (NEW YORK 100 MILES) A priori—pretrip knowledge from maps, etc.

[a]Applicable to "service-macro" destinations.

macro-performance activity of finding available emergency services. A situational example is traffic congestion that could lead to the macro-performance activity of finding an alternate route. The manner in which the driver accomplishes the in-transit phase of the macro-performance level is essentially cognitive. He searches for, or has his attention drawn to, macro-performance information, which he compares with his trip plan to decide what control action is required.

In-trip presentation of macro-performance information is made primarily by means of guide and service signs. However, receipt of information from in-trip sources other than signs (landmarks, service stations, billboards, etc.) is also a factor.

Because information received from guide and service signs is verbal or symbolic the cognitive level required for macro-performance behavior is almost entirely verbal and abstract and may required complex decisions on the part of the driver.

Inventory of Information Needs

Information needs were categorized in accordance with information inputs to the driver. The results of this categorization were combined with the levels of performance to provide eight categories of information need: (a) vehicle micro-performance; (b) ARI micro-performance (ARI refers to the advisory, restrictive, or inhibitory factors that cannot be specifically categorized under vehicle, road, traffic, service, or directional); (c) road micro-situational performance; (d) traffic situational; (e) ARI situational; (f) service macro-performance; (g) directional macro-performance; and (h) ARI macro-performance.

Table 2 is an example of one table from the inventory. A full set of tables is given in the project report (3).

DISCUSSION OF FINDINGS

This paper has characterized the driving task and information needed by the driver. The practical utility of the findings is their application in driver information systems. This discussion, therefore, considers the design of information systems for the driver.

Basic Requirements

Based on this conceptualization of the driving task, an inventory of drivers' information needs arrived at through task analysis procedures, and considerations of the nature of the driving population and of the highway system that have been detailed elsewhere (3), a set of basic requirements of a highway information system was derived. The system must be (a) user centered; (b) applicable to the existing highway system; (c) usable by all drivers at all times; (d) fail-safe; (e) compatible and evolutionary; and (f) economically feasible.

All elements and interfaces of the system should be evaluated with respect to these requirements and, in the case of conflicts, a minimax solution should be found.

Application of Factors

Five basic tenets for the systematic presentation of information needed by the driver are as follows:

1. First things first—primacy;
2. Do not overload—processing channel limitations;
3. Do it before they get on the road—a priori knowledge;
4. Keep them busy—spreading; and
5. Do not surprise them—expectancy.

Primacy—Driver information needs must be satisfied in accordance with the objective primacy of the highway system. Because of the way in which the driving task is performed and the differences in driver behavior at each level of performance, the form of information presentation at each level will differ.

Because control (micro-performance) information is highest on the primacy scale, it must be presented before guidance (situational) and navigation (macro-performance) information whenever competing needs exist. It is suggested that the information providing potential of the vehicle subsystem be optimized by appropriate design.

Two means are recommended for satisfying road micro-situational performance information needs. The first is to provide continuous adequate marking and delineations so that the driver can determine immediately his lateral and longitudinal position on the road. The use of some means of telling the driver when he is running off the road, or inadvertently changing lanes, is also recommended. This aspect of the micro-performance is amenable to in-vehicle display. The second means is by eliminating these needs by design. Avoidance of severe alignment changes, poor road surfaces, and difficult grade changes minimizes the driver's need for road micro-situational information.

Although the information needs associated with the situational level of performance are lower on the primacy scale than the micro-performance, the importance of satisfying them adequately should not be understated. Once micro-performance information needs are satisfied, situational performance needs have highest primacy.

The driver is required to rely on his capability for estimation, prediction, and judgment to perform adequately at the situational performance level. In high traffic density situations, because of the more complex decisions required, there is a higher probability that the driver will make a mistake in estimation, prediction, or judgment. Because errors at this level of performance can and usually do have catastrophic results, it is recommended that a maximum application of formal aiding techniques be made in this area. Intervehicular communication techniques should also be considered in great detail.

Macro-performance information needs are lowest on the primacy scale and are most amenable to delay. There is no way to satisfy the macro-performance needs for all drivers without considering the importance of the pretrip a priori knowledge requirements. A great portion of the macro-performance information needs can and should be satisfied before the driver even starts driving. Given adequate trip planning, all that is needed in the way of on-line directional macro-performance information presentation is information that relates the driver's trip plan to what he receives in transit—that is, telling him where he is and which way he is going. At the macro-performance level signs are a valid means of presenting information.

Processing Limitations—As the information challenge increases, a point is reached where the driver is unable to handle and process the amount of information required to resolve the uncertainty of the decision. It is at this point that the driver's channel capacity is said to have been exceeded.

As time pressure increases, it is simpler for the driver to make a series of uncomplicated decisions compared with his having to make a few more complex decisions. This is because a simpler decision takes less time; hence, more simple decisions can be made in any given time period. However, with no time pressure, the opposite is true because a driver can resolve more uncertainty in a more complex decision.

The systematic presentation of information needed by the driver must consider the processing channel limitation of the driver. The driver should not be overloaded in terms of either his attention-paying and load-shedding ability or his information-processing channel capacity.

A Priori Knowledge—The driver brings a body of knowledge, experience, and skills to the driving task. This "a priori information" is supplemented by the information acquired in preparation for a specific trip.

It can be assumed that he can add, subtract, deal in fractions, and read and comprehend simple English. It can further be assumed that the highway user starts off with the ability to operate a motor vehicle, the basic knowledge of laws and rules necessary to obtain a driver's license, and more or less specific information about his trip destination. The degree of knowledge and ability in other fields, examples of which are listed, is more uncertain and should be investigated:

1. General knowledge of geography. Distance and direction relationship of destination to origin and of destination to nearest prominent landmark or town.

2. Ability to read a map. Knowledge of where to obtain maps, and knowledge of which maps to obtain.
3. Ability to understand compass direction. Ability to translate changes of course into driving maneuvers (for example, westbound to northbound requires a right turn).
4. Ability to understand weather reports and translate them into roadway and visibility conditions.
5. Ability to translate distance into driving time under prevailing conditions.
6. Degree of familiarity with highway and interchange types and elements.

Spreading—There are times during the driving task when the driver's processing capacity is fully used or overloaded as well as times when his processing capacity is almost completely unused. Both unused processing capacity and overloaded processing capacity present serious problems.

When the driver has little to do, representative of the condition where the driver's processing capacity is almost completely unused, there are two possible problems. The first problem involves the self-pacing attribute of the driver. Self-pacing refers to the fact that drivers seem to prefer to set their pace in accordance with their own subjective concept of what they can do.

In locations where little in the way of events or signals is occurring (for example, on rural freeways), the external pacing of the road is very low. When the external pacing of the road is low, drivers have been found to create their own work in order to satisfy their self-pacing. Thus, if little is occurring, drivers may execute unnecessary and even dangerous maneuvers to do. A second problem in low signal areas is that of vigilance, where operators have been found to miss needed signals for no apparent reason. This is an important consideration for the older driver who has been found to be less vigilant than the younger driver.

In high-signal areas, where many signals are competing for the driver's attention, entirely different problems exist. If the driver's processing capacity is overloaded, he may miss signals because he was unable to load-shed properly. He may also be faced with the problem of not having enough time to make a decision, leading to confusion or decisions made on the basis of incomplete information.

One way to avoid the problem caused by too few or by too many signals is by applying the principle of spreading. In applying this principle of spreading, the driver attention demand situation with peaks (too much information) and valleys (too little information) is determined. The objective primacy of the situation is determined, and the less important information needs at the peak where too much is occurring are transferred to the valleys, thus approaching an even distribution of information presentation and minimizing the problems of low and high signal areas.

Expectancy—Expectancies are a function of the driver's experience and a priori knowledge. An important part of driver education is the development of a set of realistic expectancies.

Micro-performance expectancies are those associated with the vehicle and its position on the road. That drivers expect their vehicles to respond to their control actions is indicative of expectancy at this level. Once position on the road has been established, the driver relies on inference and prediction to maintain his position on the road. Anything occurring that is not within the realm of what the driver expects can lead to trouble. Micro-performance expectancies operate below the level of consciousness.

The importance of expectancies is perhaps the greatest at the situational level. Alignment changes provide examples of how expectancies at this level can be structured. Drivers expect some alignment changes on all roads and, depending on their experience on the particular type of road, expect to make speed and tracking corrections. Warning signs must be used to structure driver expectancies. Few drivers expect lane drops and other changes in cross section. Therefore, warnings of these changes are necessary.

Another class of expectancies involves interchange geometrics such as whether exists are on the left or right. Since left exits and entrances are seldom encountered, they are counter to driver expectancies. Currently very little information on type and configuration of interchange is presented to the driver. His expectancies are rarely structured except by a priori knowledge.

If a car's left-turn signal is on, the car is expected to turn left. Similarly, when approaching a green light, the driver expects cross traffic to stop. Intervehicular signaling is indicative of expectancy structuring at this level.

A third category of expectancies involves direction-finding information needs. An important aspect of these expectancies is that they are most amenable to structuring because so much of the macro-performance level involves pretrip preparation. The driver expects to find in-trip cues that correspond to his trip plan. He expects to find signs telling him where he is and which way he is going. He also expects to find information telling him where services are located and how to get to them. At the macro-performance level of driving, the greatest potential exists for structuring expectancies (through maps) and then satisfying these expectancies (through signs).

REFERENCES

1. Cumming, R. W. Progress Report of Human Factors Committee. Australian Road Research, Dec. 1962.
2. NCHRP Project 3-12. Project Statement, April 1966.
3. Alexander, King, and Warskow. Development of Information Requirements and Transmission Techniques for Highway Users. AIL, Division of Cutler-Hammer, Nov. 1967.
4. Greenshields, B. D. Driving Behavior and Related Problems. Highway Research Record 35, 1963, pp. 14-31.
5. Platt, F. M. A New Method of Measuring the Effects of Continued Driving Performance. Highway Research Record 25, 1963, pp. 33-57.
6. Algea, C. W. A Development of a Conceptual Framework for the Driving Task. Human Factors, Vol. 6, No. 4, 1964, pp. 375-382.
7. Biggs, N. L. Directional Guidance of Motor Vehicles—A Preliminary Survey and Analysis. Ergonomics, Vol. 9, No. 3, 1966, pp. 193-202.
8. Miller, R. B. A Method for Man-Machine Task Analysis. WADC-TR-53-137, Wright Air Development Center, Dayton, Ohio, 1953.
9. Seale, L. M. Purpose of Task Analysis, Methods of Obtaining Task Analysis Information, Time Phasing. Conference on the Uses of Task Analysis in the Derivation of Training and Training Equipment, Wright Air Development Center, Dayton, Ohio, 1959.
10. Alexander, G. J. Task Analysis to Determine Information Needs and Control Actions for Personnel in a Military Tactical Mission. AIL, Division of Cutler-Hammer, 1966 (unpublished).
11. Rashevsky, N. Man-Machine Interaction in Automobile Driving. Traffic Safety Research Review, Vol. 9, 1965, pp. 101-107.
12. Gordon, D. A. Experimental Isolation of the Driver's Visual Input. Human Factors, Vol. 8, No. 2, 1966, pp. 129-137.
13. Michaels, R. M., and Cozan, L. W. Perceptual and Field Factors Causing Lateral Displacement. Highway Research Record 25, 1963, pp. 1-13.
14. Lefeve, B. A. Speed Characteristics on Vertical Curves. HRB Proc., Vol. 32, 1953, pp. 395-413.
15. Barch, A. M. Judgments of Speed on the Open Highway. Jour. Appl. Psychol., Vol. 42, No. 6, 1958, pp. 362-366.
16. Denton, G. G. A Subjective Scale of Speed When Driving a Motor Vehicle. Ergonomics, Vol. 9, No. 3, 1966, pp. 203-210.
17. Todosiev, E. P. The Action Point Model of the Driver-Vehicle System. Engineering Experiment Station Report EES 202A-3, Ohio State University, 1963.
18. Wohl, J. G. Man-Machine Steering Dynamics. Human Factors, Vol. 3, No. 4, 1961, pp. 222-228.
19. Cumming, R. W. The Analysis of Skills in Driving. Australian Road Research, Vol. 9, 1964, pp. 4-14.
20. Schlesinger, L. E., and Safren, M. A. Perceptual Analysis of the Driving Task. Highway Research Record 84, 1963, pp. 54-61.

Olson, P. L., Wachsler, R. A., and Bauer, H. J. Driver Judgments of Relative Car Velocities. Jour. Appl. Psychol., Vol. 45, No. 3, 1961, pp. 161-164.

Braunstein, M. L., and Laughery, K. R. Detection of Vehicle Velocity Changes During Expressway Driving. Human Factors, Vol. 6, No. 4, 1964, pp. 327-331.

Hoppe, D. R., and Lauer, A. R. Factors Affecting the Perception of Relative Motion and Distance Between Vehicles at Night. HRB Bull. 43, 1951, pp. 1-17.

Stalder, H. I., and Lauer, A. R. Effect of Pattern Distribution on Perception of Relative Motion in Low Levels of Illumination. HRB Bull. 56, 1952, pp. 25-36.

Wright, S., and Sleight, R. B. How Close Is Too Close? Traffic Safety, 1961.

Lerner, H. D., Abbott, H. E., and Sleight, R. B. Following Distance on the Highway Related to Speed, Trip Duration, Traffic and Illumination. Human Factors, Vol. 6, No. 4, 1964, pp. 343-350.

Bierley, R. L. Investigation of an Intervehicle Spacing Display. Highway Research Record 25, 1963, pp. 58-75.

Fenton, R. E. An Improved Man-Machine Interface for the Driver-Vehicle System. Report given at the 7th IEEE Symposium on Human Factors in Electronics, May 1966.

Bjorkman, M. An Exploratory Study of Predictive Judgments in a Traffic Situation. Scandinavian Jour. of Psychol., Vol. 4, No. 2, 1962, pp. 65-76.

Crawford, A. The Overtaking Driver. Ergonomics, Vol. 6, No. 2, 1963, pp. 153-170.

Jones, H. V., and Heimstra, N. W. Ability of Drivers to Make Critical Passing Judgments. Jour. Eng. Psychol. Vol. 3, No. 4, 1964, pp. 117-122.

Brown, I. D. Effect of a Car Radio on Driving in Traffic. Ergonomics, Vol. 8, No. 4, 1965, pp. 475-479.

Hulbert, S. F. Drivers' GSR's in Traffic. Percept. Mot. Skills, Vol. 7, 1957, pp. 305-315.

Taylor, D. H. Drivers' Galvanic Skin Response and the Risk of Accident. Ergonomics, Vol. 7, No. 4, 1964, pp. 439-451.

Gibbs, C. B., Davidson, A., and Shackel, B. Car Turning Signals and Delays in Responding to Visual Information. APU 176/52, Sept. 1952.

Shor, R. E. Shared Patterns of Nonverbal Normative Expectations in Automobile Driving. Jour. Soc. Psychol., Vol. 62 (First Half), 1964, pp. 155-163.

A COMPARISON OF SELECTED TRAFFIC INFORMATION DEVICES

Gerald C. Hoff, Chicago Area Expressway Surveillance Project,
Illinois Division of Highways

This study evaluates various traffic information devices with respect to their preference rating by drivers and to the response to each device. A total of 187 respondents were interviewed in groups ranging from 6 to 55. The 6 alternates that were offered to the respondents for preference rating were a symbolic expressway map sign with colored arrows showing traffic conditions, symbolic expressway and arterial street map sign with colored arrows showing traffic conditions, changeable-message lamp matrix sign, roadside radio transmitter, commercial radio traffic broadcast, and experience and driving background. The respondents preferred the lamp matrix sign, the symbolic maps, the radio system, and experience, in that order. The respondents were asked to make a diversion decision based on traffic information received from each of the devices. Information concerning 6 different levels of congestion was presented on the lamp matrix sign, the symbolic map, the roadside radio, and the commercial radio traffic broadcast. For each device, diversion increased as congestion increased, except that the diversion proportion for the symbolic map at one congestion level deviated sharply from an otherwise logical trend. The results show that the respondents preferred visual devices over vocal devices and, for the sample, the differences in diversion response between devices were significant but small in a practical sense.

•DURING the past few years, the motoring public has been led to believe that the major breakthrough necessary to automate their driving is just around the corner. The engineers responsible for operating the highway system would like to believe that this is true. However, it appears that a more realistic time schedule will include a lengthy research program followed by an equally lengthy implementation program. This time schedule must include not only all the necessary research and hardware production time but also the effort necessary to motivate the drivers to buy and use this improvement.

Although current research concerning automatic control of vehicles has produced some significant results, much more work remains to be completed before a safe, economical, and operationally satisfactory control system can be mass produced (1). The automobile driver in the United States has never enthusiastically relinquished any of the independence he enjoys in operating his vehicle. For example, he continues to drive in urban areas where in many cases the trip could be completed more economically and with little time differential by mass transit. To be successful, the automated highway will require a wholesale change in values of drivers.

BACKGROUND

While waiting for the potentially more efficient automated highway to become a reality, the highway engineer has coped with the astounding growth in traffic by imple-

Sponsored by Committee on Motorist Information Systems and presented at the 50th Annual Meeting.

menting changes in the highway system. Some of these changes have increased the highway system's capacity by providing more and better designed roadways. Other changes have attempted to utilize existing roadways more efficiently. This class of changes has included more sophisticated traffic signal systems, better signing, freeway surveillance and control systems, and experimentation in electronic route guidance systems.

One potential improvement, furnishing the motorist with accurate and up-to-date traffic information, has been the subject of only a limited amount of study. Assuming that travel time plays an important part in the driver's selection of one route from a number of alternates, timely and accurate traffic information may help the driver to select the route that not only has an acceptable travel time but also improves the efficiency of the highway system.

The implementation of any type of transportation improvement should be predicated on affirmative answers to three questions:

1. Will this improvement satisfy a current or projected need?
2. Is the technology capable of supporting this improvement?
3. At the indicated level of expenditure, do the positive results of this improvement outweigh any negative results?

Is There a Need?

The driver's awareness of the need for more or better information to assist him in his driving task has been demonstrated in a number of earlier studies. Heathington tested 782 Chicago area drivers to determine what priorities they gave to improvements that could be made to urban freeways (2). The top three choices were:

1. Better repair of pavement damages such as holes and bumps;
2. Increased enforcement of regulations concerning shoulder riding, lane changing, and driving speed (minimum and maximum); and
3. Provision of signs that can be electronically changed to furnish information about traffic conditions on the expressway ahead.

In the study done by Covault and Bowes on the Kentucky Toll Road, motorists were given information by radio concerning conditions on the roadway ahead (3). A large majority (90 percent) considered the combination of radio and signs preferable to signs alone, particularly when unusual conditions—accidents, maintenance operations, inclement weather, detours, etc.—were encountered. Of the sample, 75 percent would pay over $15 for such a radio, 48 percent over $30, and 25 percent over $50. Only 8 percent would not purchase such a radio.

A recent inventory of attitudes toward transportation by McMillan and Assael found that 62 percent of their sample would like to see the same amount or more money spent on improved traffic signals and signing (4). This ranked fourth among eleven suggested transportation improvements.

One of the outputs of the experimentation in electronic route guidance was a market research study conducted by Eberhard (5). The study compared an in-car route guidance display with conventional signing. A very high proportion (85 percent) of the respondents thought such a system better than the conventional signs. The cost of such a system was estimated at $95 by the respondents. They felt primary use would occur on highways connecting suburban areas with the central business district.

These studies confirm the suspicion that drivers are somewhat dissatisfied with the current method of providing both traffic and routing information. Even recognizing that there is some personal cost included, most drivers prefer to receive a higher level of information than they are now getting.

Most large cities have some form of radio traffic broadcasts, usually consisting of a combination of information from a traffic center and from mobile observers, either airborne or in cars. Because almost all of these broadcasts are commercially sponsored, it would follow that the sponsors consider their expenditure to be offset by the goodwill accruing to them as a result of the broadcasts. In fact, in the Chicago area, some of the broadcasts are sponsored by the commuter railroads. The interest of the sponsors may be considered to be a reflection of the motorists' desire for more traffic information.

Those studies conducted to date and the apparent commercial success of radio traffic broadcasts point out an apparent deficiency of traffic information from the motorist's viewpoint.

Can It Be Done?

The collection of the base of information necessary to support a real-time traffic information system is well within the ability of current technology. As traffic control systems have become more sophisticated, the amount of traffic information collected at a central control point has increased.

An example of this is the Wichita Falls, Texas, computerized traffic control installation (6). The system uses a digital computer to control 77 signalized intersections. The primary objective is to reduce delay and stops at these intersections. The control scheme is based on data acquired from 51 vehicle detectors at 26 different locations. For each detector, the computer monitors (a) the number of vehicles detected; (b) the number of vehicles required to stop; (c) the total stopped delay (vehicle-seconds); (d) the average delay per vehicle; and (e) the probability that a vehicle will be stopped. This type of digital computer-controlled traffic signal system will become more popular as traffic engineers gain familiarity with the resulting increase in capability and flexibility.

As with computerized traffic control systems, the use of freeway surveillance and ramp control systems is becoming more prevalent. Current operational installations in Chicago, Detroit, and Houston will soon be joined by systems in Los Angeles, Boston, and other large cities. Because the successful operation of such a system is directly dependent on the collection and analysis of freeway traffic flow data, this information can become available for motorists to use to select their route (7).

The analysis of the data is normally carried out by a digital computer. When properly programmed, the computer without any intervention can extract the pertinent information from the data and prepare it for presentation to the public with little if any time delay.

The limits of technology do not pose a critical problem in determining how the information is presented. Many presentation techniques are technologically sound and could be implemented at once. The information could be presented to the driver either in his car or outside of his car and could be either oral or visual. Although oral forms will probably be restricted to radio broadcasts of some type, visual forms can be signs outside of the car or a heads-up display inside the car (8).

Although it has so far been a matter of personal opinion and conjecture as to which technique might be the most effective, the ability of the technology to support both data collection and information presentation cannot be a serious deterrent to further prosecution of the development of a traffic information system.

What Might Such a System Do?

In reviewing past research into the effect that an increased level of information has on traffic movement, one should recall that each study was of a particular hardware configuration. The results of these studies reflect the reaction of the driver to the hardware used as well as his reaction to the information presented by the hardware.

With some possible exceptions, no study has directly addressed the problem of the effect of traffic information on route selection. One exception was an inconclusive evaluation by the author of the effect of color-coded changeable-message displays on selection of the best ramp at which to enter an expressway (9). In another study, Heathington found that drivers diverted from their normal route 23 percent of the time when the traffic was heavy (2). The diversion increases somewhat to 30 percent if the driver has heard a radio report of an accident on his normal route.

Other studies have measured the relationship between driver response and information through the use of such parameters as speeds and lane changing. Covault and Bowes found that a roadside radio message stating that an accident had occurred ahead on the left caused a speed decrease of 5 to 8 mph and significant shift of traffic to the right (3). Other messages concerning shoulder maintenance and grass cutting caused a decrease in speed of up to 10 mph but no significant change in lateral placement.

The use of roadside radio as a route guidance technique has also been studied (10). The additional information provided caused the driver to enter the deceleration lane of a designated exit ramp at an earlier point and to make the diverging maneuver at a higher speed.

The use of lane control signals and advisory speed signs on the Lodge Expressway were examined by Wattleworth and Wallace (11). The study considered the use of these devices during both peak and off-peak conditions. During the off-peak, the lane control signals had some significant but small benefits in total travel time and lateral placement. The use of the changeable speed-limit signs during the off-peak produced negligible benefits. The use of these devices during the peak period was judged to have produced essentially no effect.

A series of blank-out signs indicating the closure of entrance ramps was tested to determine their ability to divert entrance ramp traffic. These signs were mounted at all approaches to the ramps under study. They were remotely controlled to indicate that the ramp was closed when traffic conditions warranted. No other device was used to close the ramps. About 25 percent of the normal entrance ramp traffic was diverted during the time these signs were on. The most diversion took place at those ramps where good alternates were available. Also, the entrance ramp motorist could see the freeway from the frontage road before he was committed to using it, and some diversion could be attributed to observation rather than to reaction to the sign.

From these studies it is evident that without further research a reliable estimate of the diversion or traffic flow improvements resulting from a real-time traffic information system could not be made.

This brief look at the previous examinations of traffic information systems has provided some answers to the questions posed for consideration of a transportation improvement. The motorist is dissatisfied with the present level and quality of traffic and routing information that he receives. In particular, information about urban freeway traffic seems to be a primary concern. The technological capability exists to provide both route and traffic information. The major decision to be made concerns the technique of presentation. Estimating the results of a driver information system cannot be done without additional study. The technique chosen to present the information will in part determine the effectiveness of the system. It is the intent of this paper to provide information to be used in making the selection of the system that offers the greatest chance of success.

CONDUCT OF THE RESEARCH

The inconclusive results of an earlier effort to divert potential entrance ramp traffic from congested expressways created a feeling that perhaps the installation of the informational sign was premature (9). A more reasonable but more lengthy course of action would have been to first examine the various techniques that might be used to present traffic information to ascertain which techniques are the most feasible and which offered an acceptable potential for success.

Motorists headed for an expressway entrance ramp can benefit from diversion under two sets of circumstances. One is when the expressway has become severely congested as a result of either an incident or an excess of ramp traffic attempting to use the expressway. Incidents are quite common, with about three peak periods out of four having reduced capacity due to an incident (12).

The other occasion when the ramp user might benefit from diversion will occur as the result of the operation of a ramp control system. Such a system can result in queuing behind ramp metering devices. If the queue delay is longer than the increased travel time on an alternate route, the motorist will benefit by diverting.

Diversion can also increase the efficiency of the roadway network. A ramp control system must cause some diversion in order to accomplish its purpose. The accumulation of travel time benefits on the expressway can be offset by ramp delays if no diversion takes place. Other delays can occur if the ramp queues extend into adjacent intersections and traffic other than that directly involved with the operation of the freeway is delayed. This delay is associated with the degraded operation of the intersection, and the amount of time lost can increase very rapidly.

In order to benefit the roadway system, diversion must be relied on to keep ramp delays from overriding expressway benefits and must keep ramp queues from interfering with arterial street traffic.

The Chicago Area Expressway Surveillance Project is rapidly expanding its system of freeway surveillance and ramp control. An increased length of freeway under surveillance and an increase in the number of controlled ramps have provided both the means and the necessity to become involved in traffic information research.

Because the project already had the data collection and analysis capability, the selection of a presentation device was the first critical decision to be made. This decision could not be made without further information about the motorist's opinion of the various devices under consideration.

Development of Survey Questionnaire

A questionnaire was developed that would determine how well a motorist liked each candidate device and how strongly and correctly he responded to each device. The questionnaire was divided into six sections. Some of the sections are not germane to this particular research but were included to provide data on some other problems of current interest.

The first section compared the driver's ability to identify expressways in Chicago by name and number. The second section tested some of the hypotheses used in the design of an earlier informational sign developed by the Chicago Area Expressway Surveillance Project (9). An explanation of how the symbols used on that sign correlated with traffic conditions constituted the third section.

Sections four and five are being used as the basis of this report. Section four of the questionnaire was designed to place the candidate devices on a preference scale. This was accomplished through two independent techniques and the results were compared. The preference scale shows whether oral or visual information is preferred and identifies the particular device liked best by the motorist.

Each device presented freeway traffic conditions and conditions at controlled ramps. Because the surveillance project is currently operating a ramp control system, the staff felt that unusual ramp delay might be important to the motorist.

The candidate devices tested were as follows:

1. A symbolic map with arrows showing traffic conditions (Fig. 1a);
2. A symbolic map with arrows and an indication of where the driver was within the highway system (Fig. 1b);
3. A changeable-message matrix sign (Fig. 1c);
4. A roadside radio system;
5. A commercial radio traffic broadcast; and
6. A null alternative, using experience only.

The purpose of the fifth portion of the questionnaire was to determine driver response to traffic information as presented on some of the devices previously listed. Devices 1, 3, 4, and 5 were each used to provide information about six levels of congestion. The respondent was asked to make a route selection decision based on the information from each device for each level of congestion.

Comparisons by levels of congestion and by device were made. By comparing the response to different devices at the same level of congestion, the strength of the devices could be examined. By comparing the response to different levels of congestion as presented on the same device, the ability of the device to promote a consistent response pattern could be probed.

Socioeconomic and travel characteristics were obtained in the sixth section.

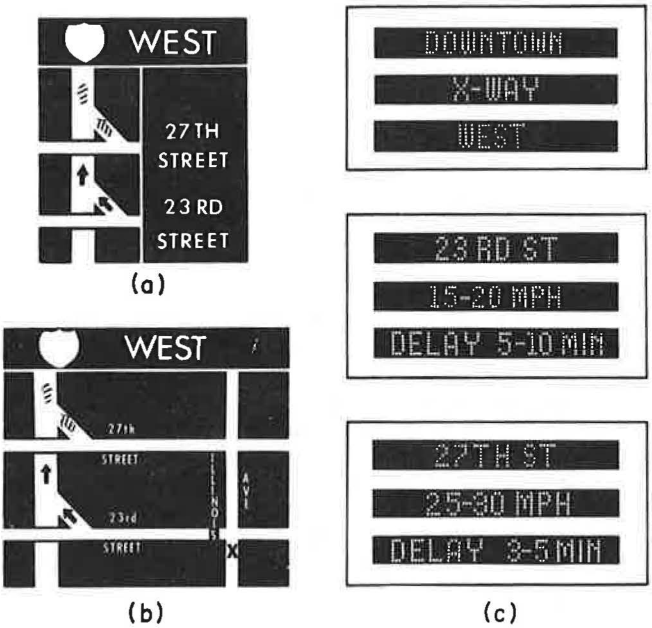


Figure 1. (a) Symbolic map (arrows change color); (b) symbolic map with arterial streets (arrows change color); (c) matrix sign.

Administration of the Questionnaire

Extensive use was made of slides, movies, and tape recordings to demonstrate each device as realistically as possible. Therefore, separate tests for each individual were not practical, and administration of the questionnaire to groups was necessary. Groups consisted of Army Reserve personnel, junior college students, truck driver trainees, and employees of Chicago's Bureau of Public Works. The groups ranged in size from 12 to 60 individuals, and 189 usable questionnaires were obtained.

DRIVER PREFERENCE

In the 1920s, Thurstone developed an experimental technique with which a set of stimuli could be placed on a scale indicating some psychological judgment such as preference, beauty, or loudness (13). The technique consists of presenting to a respondent a pair of stimuli—in this case two different information presentation devices—and asking him to choose the one that best meets some stated criteria.

All possible paired combinations of stimuli are presented for choice. If n stimuli are being considered and there would be no difference in response due to order within the pair, $(n(n-1))/2$ pairs must be tested. By having one respondent make the choice of the same stimulus pairs many times or by having a large number of respondents make the choice once, a frequency of choice of one stimulus over another stimulus can be derived. The frequency of preference of stimulus j over stimulus k for all pairs of stimuli can then be summarized in tabular form.

One can picture the discriminantal process associated with evaluation of a stimulus on a psychological continuum as in Figure 2. The scale value given to the stimulus is S_j, which is the mean of the distribution of the discriminantal process on the continuum. The distribution is not single-valued because at different times or under different conditions variations can be expected to occur in the discriminantal process. Thurstone has assumed that this distribution is normal and can be adequately described by the mean and variance.

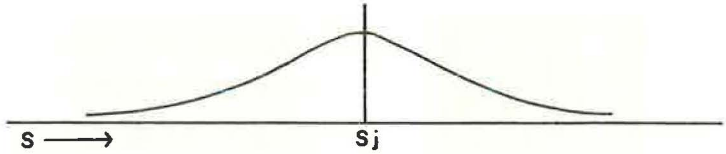


Figure 2. Discriminantal process.

When two stimuli are being compared, the distance between the two means is used to provide the scale distance along the continuum (Fig. 3).

Any single presentation of a stimulus to an observer results in a discriminantal process: d_j or d_k. The difference in the discriminantal processes $(d_k - d_j)$ is termed a discriminantal difference. This discriminantal difference also forms a normal distribution on a continuum. The mean value of the $d_k - d_j$ distribution is used as the difference in scale values for the two stimuli. The distribution of discriminantal differences in shown in Figure 4.

The shaded portion shows the proportion of the distribution where $(d_k - d_j)$ is positive. The value X_{kj} is the mean and is measured in units of the deviation, $\sigma_{d_k - d_j}$ of the distribution. Then

$$S_k - S_j = X_{jk} \cdot \sigma_{d_k - d_j}$$

since

$$\sigma_{d_k - d_j} = (\sigma_j^2 + \sigma_k^2 - 2r\sigma_j\sigma_k)^{1/2}$$

where r = correlation between σ_j and σ_k

$$S_k - S_j = X_{jk} (\sigma_j^2 + \sigma_k^2 - 2r\sigma_j\sigma_k)^{1/2}$$

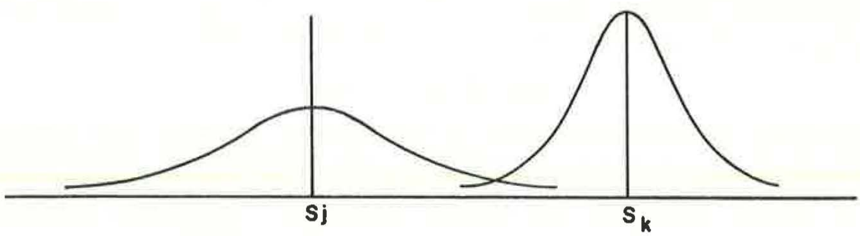


Figure 3. Comparison of two processes.

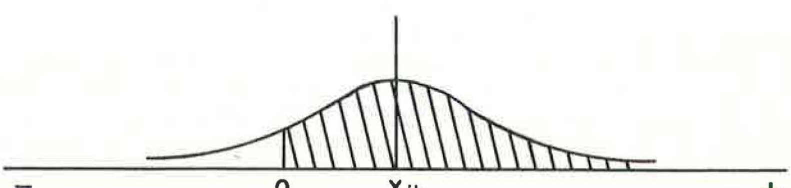


Figure 4. Distribution of discriminantal differences.

which is the formal statement of Thurstone's Law of Comparative Judgment.

The equation in this form is insoluble because there are more unknowns than available equations. Various simplifying assumptions have been worked out by both Thurstone and Torgerson (14, 15). The pertinent simplification to be used here is Thurstone's Case 5, which assumes that the deviations are equal and the correlation is zero for each pair of stimuli. The law then reduces to

$$S_k - S_j = X_{jk}C$$

X_{jk} = discriminantal difference

Torgerson notes that the same solution can be obtained by assuming equal deviations and equal correlations. Since C is only a scale factor it is set equal to unity.

Calculation of Preference Scale

A series of respondents are required to make a preference judgment for each of the stimuli pairs. From their responses a matrix indicating the proportion of trials that j is preferred to k is prepared. It should be noted that

$$P_{kj} = I - P_{jk}$$

P_{kj} = Proportion of trials in which k is preferred to j
P_{jk} = Proportion of trials in which j is preferred to k

From this proportion matrix, another matrix, the X matrix, is constructed. Each cell of the X matrix is equal to the standard normal deviate corresponding to the same cell in the proportion matrix. The scale values are then obtained by averaging the columns of the X matrix. Usually the smallest scale value is assumed to be the zero point and the remaining values are adjusted accordingly.

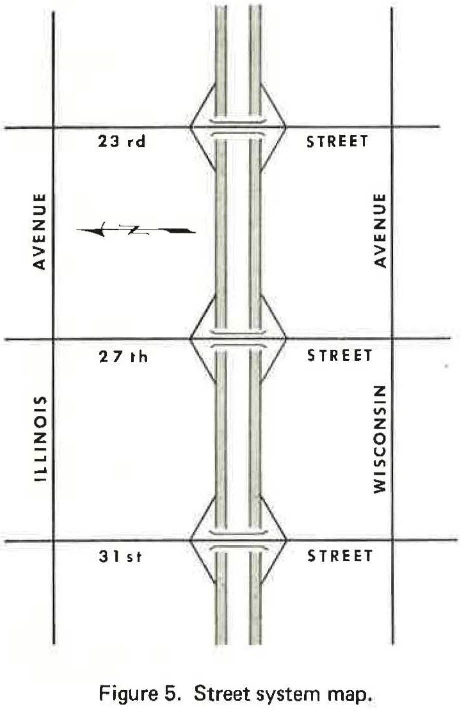


Figure 5. Street system map.

Conduct of Experiment

Five techniques for presenting traffic information to the driver were used as stimuli along with an alternative representing the current method based on driving experience used as a null alternative. The respondents were first introduced to the concept of color coding of traffic information in the questionnaire through the use of a short motion picture. The film showed traffic conditions on an expressway from both an overhead and an in-car view. The first portion of the film showed the traffic conditions corresponding to the congestion level at which the freeway arrows on a symbolic map (Figs. 1a and 1b) would change from green to yellow. The next portion of the film presented the traffic conditions on the expressway which would cause the arrow to go from yellow to red. Finally, the process of ramp metering was shown and the problem of queue development was explained. In this explanation, the inability of current detection systems to estimate ramp delay accurately was mentioned.

The following traffic situation was explained with the use of a map (Fig. 5):

"Suppose that on the expressway westbound traffic is moving at 20 mph at 23rd Street, 25 to 30 mph at 27th Street, and 40 to 45 mph at 31st Street. You want to use the westbound Downtown Expressway and hope to enter at one of the ramps shown on the map. All of these ramps are controlled. There is a 5- to 10-minute delay at the 23rd Street ramp, about a 3-minute delay at the 27th Street ramp, and less than a minute delay at the 31st Street ramp."

The manner in which each of the six alternative techniques would present the information from this situation was demonstrated to the respondents:

Alternative 1, Driving Experience—A slide depicting a suburban arterial intersection was shown.

Alternative 2, Symbolic Map With Arrows—A slide of the device was shown with the freeway and ramp arrows for 23rd Street red and the arrows for 27th Street yellow.

Alternative 3, Symbolic Map With Arrows and Arterial Street—The same information was shown as in alternative 2.

Alternative 4, Changeable Message Matrix Sign—Three successive messages were shown. Each message was shown for 2 sec to simulate viewing time while approaching the sign, as follows: First,

DOWNTOWN
X-WAY
WEST

Second,

23RD ST
15-20 MPH
DELAY 5-10 MIN

Third,

27TH ST
25-30 MPH
DELAY 3-5 MIN

Alternative 5, Roadside Radio—A cartoon of a car with a small radio transmitter adjacent to it was shown to accompany the following tape recording:

"At the 23rd Street entrance ramp to the westbound Downtown Expressway, there is a 5- to 10-minute delay and the expressway is moving at 15 to 20 miles an hour. At the 27th Street entrance ramp there is a 3- to 5-minute delay and expressway speeds are 25 to 30 miles an hour. At 31st Street there is no ramp delay and speeds are 45 miles an hour. To reach 27th or 31st Street, take a right at the next intersection."

Alternative 6, Commercial Radio Traffic Broadcast—A cartoon of a helicopter transmitting information to a car was used to illustrate the following tape recording:

"And now on the westbound Downtown Expressway, traffic is heavy to 29th Street where it opens up. There is a ½-block back-up at the 23rd Street entrance ramp."

The information presented by each alternative technique was made to correspond as closely as possible to the capabilities of the technique and the data that would be available for its use.

The respondents were then presented all possible pair combinations of the six alternatives and asked to indicate on the response form which member of the pair they felt gave the best information. The order in which the pairs and the members within the pairs were presented was randomized to ensure that the same sequence was not used for any test group.

Results of the Experiment

The 189 questionnaires were summarized and the proportion matrix was constructed (Table 1). From these data the X matrix was constructed and the scale was derived as given in Table 2. A scale diagram was made as a graphic presentation of the results (Fig. 6).

TABLE 1

PREFERENCE PROPORTION

Alternative k	Alternative j					
	1	2	3	4	5	6
1	—	0.8836	0.8730	0.8836	0.7619	0.7937
2	0.1164	—	0.5714	0.6243	0.4286	0.4021
3	0.1270	0.4286	—	0.5926	0.3651	0.4392
4	0.1164	0.3757	0.4074	—	0.3122	0.3492
5	0.2381	0.5714	0.6349	0.6878	—	0.5608
6	0.2063	0.5979	0.3608	0.6508	0.4392	—

Note: Each cell represents the proportion of trials that alternative j was preferred to alternative k.

Alternatives:

1. Experience
2. Symbolic map with arrows
3. Symbolic map with arrows and arterial streets
4. Changeable message matrix
5. Roadside radio
6. Commercial radio

TABLE 2

X MATRIX

Alternative k	Alternative j					
	1	2	3	4	5	6
1	—	1.19	1.14	1.19	0.71	0.81
2	-1.19	—	0.18	0.31	-0.18	-0.24
3	-1.14	-0.18	—	0.23	-0.34	-0.15
4	-1.19	-0.31	-0.23	—	-0.49	-0.38
5	-0.71	0.18	0.34	0.49	—	0.15
6	-0.81	0.24	0.15	0.38	-0.15	—
Xka	-5.04	1.12	1.58	2.60	-0.45	0.19
Xka/n	0.84	0.19	0.26	0.43	-0.08	0.03
Scale	0.00	1.03	1.10	1.27	0.76	0.87

In general, visual sources of information appeared to be preferred to oral sources. As could be expected, experience was the least preferred alternative. Using a method suggested by Heathington, successive pairs of alternatives were tested for significant difference (2). All were significant at the 95 percent level except the two symbolic maps and the two radio techniques. The internal consistency of the data was checked and found to be adequate.

In summary, then, if a driver were able to choose a device for presentation of traffic information, a changeable message matrix sign would be preferred. The choice is not unexpected when the results of Heathington's work, which showed that drivers prefer speed to any other traffic descriptor, are taken into consideration (2). Because the symbolic maps with color-coded arrows provide a qualitative description of traffic conditions, they would be less preferred than the changeable message matrix, which provides speed data.

The general preference for visual displays over oral presentations, although not necessarily surprising, is more difficult to rationalize. Perhaps the driver feels that he may have a difficult time extracting the pertinent information from an oral presentation. If he is not listening very carefully, he may miss the information that is important to him.

The changeable message matrix sign offers one important advantage over the other visual techniques. With the proper control circuitry, the information displayed can be changed to provide specific information about special traffic situations such as accidents, lane blockages, and reduced capacity.

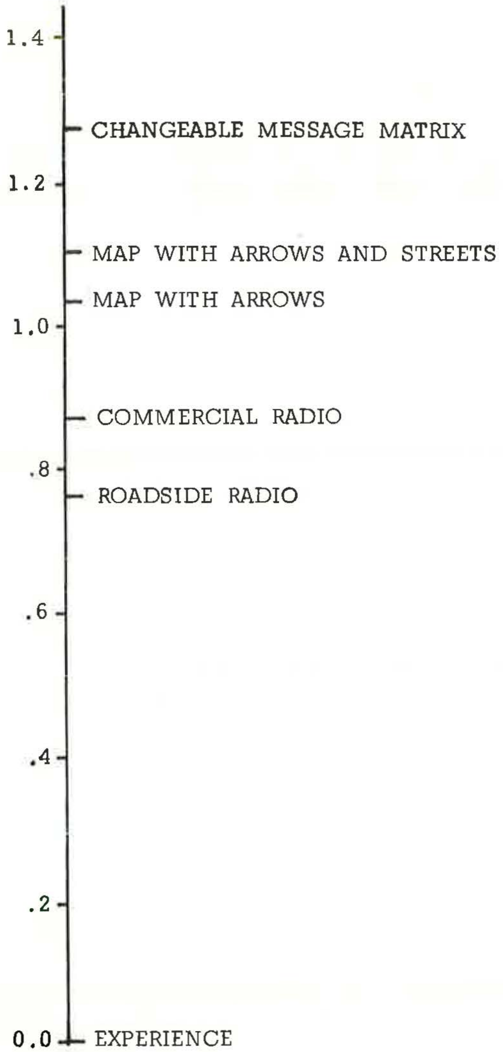


Figure 6. Preference scale.

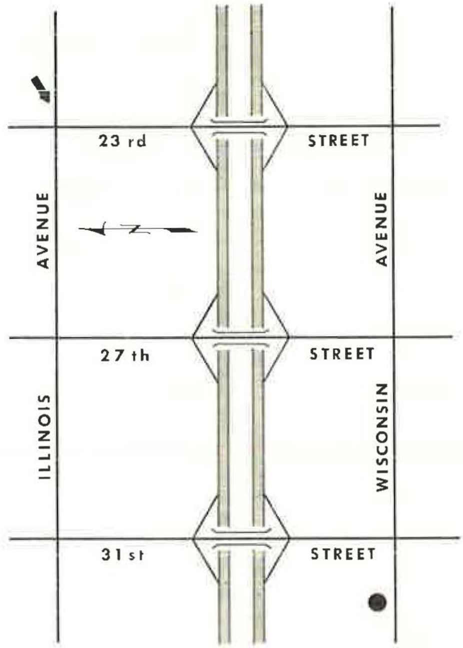


Figure 7. Diversion map.

DRIVER RESPONSE

Ultimately, a traffic information system must be able to influence the actions of drivers in such a way that they are encouraged to increase the level of efficiency in the roadway system. The device most preferred by the driver for presenting information will not necessarily cause the proper actions on the part of the driver to occur. It is therefore necessary to investigate how the driver responds to each of the candidate systems in terms of whether he makes a correct route selection based on the information that is given to him, and which device provokes the strongest response based on a given level of traffic information.

In order to test driver response, each respondent was presented with a traffic situation in which he had to decide whether to use an expressway route or an alternate route. He was given traffic information concerning the expressway route, and this information was varied to provide six levels of traffic service on the expressway. Each of these six levels was presented over four of the candidate systems.

The first candidate system used was the symbolic map with colored arrows. The second candidate system used was the changeable message matrix sign. The third system was the roadside radio, and commercial traffic broadcast was the fourth system considered.

The questionnaire introduced the traffic situation as follows:

"Now I'm going to ask you to pretend that you are driving a car on your way home. It is the evening rush hour, and you are planning to use the expressway. You are approaching an intersection where you must decide whether to use the nearest ramp to get on the expressway, or to use the city streets to go to the next nearest ramp, or to use the city streets for your entire journey. The situation is shown on this map (Fig. 7). Your car is in the same position as the arrow. Your home is about 10 miles away, shown by the circle. The city street is an average suburban business street. The traffic conditions are shown in this film. (The film depicted a 3-minute trip down the suburban arterial street in the Chicago metropolitan area.) Now I am going to give you some traffic information and ask you if you will (a) use the nearest ramp, or (b) not use the nearest ramp."

The information was presented in the same format that was used in the preference scaling part of the questionnaire; however, in this case the level of congestion in the area of 23rd Street was varied. The level of congestion on other portions of the expressway remained the same. Six levels of congestion were chosen at 23rd Street to be representative of most of the conditions that would be considered by a motorist. The manner in which the information was presented is given in Table 3.

For each group the order in which the information was presented was randomized to minimize the effects that might have occurred because of the people remembering how they responded to a particular piece of information as presented by another device or how they responded to other information presented by the same device. The information was summarized as the proportion of the sample that indicated they would divert from the nearest entrance ramp based on the information as it was presented to them. This information is shown in Figure 8.

Comparison of Devices

The diversion patterns shown in Figure 8 were used to compare the devices against each other to determine whether one device would cause a significantly higher or lower

TABLE 3

CONGESTION LEVEL INFORMATION

Amount of Delay Caused by Using 23rd St. Ramp	Symbolic Map[a]	Changeable Message Matrix[b]	Roadside Radio[c]	Commercial Traffic[d]
-2 min	Freeway-Green Ramp-Green	23rd St 45-55 mph Delay 0-3 min	A. 0 to 3 B. 45 to 55	A. Light B. No
+1 min	Freeway-Green Ramp-Yellow	23rd St 45-55 mph Delay 3-5 min	A. 3 to 5 B. 45 to 55	A. Light B. No
+3 min	Freeway-Yellow Ramp-Yellow	23rd St 25-30 mph Delay 3-5 min	A. 3 to 5 B. 25 to 30	A. Moderate B. A short
+5 min	Freeway-Red Ramp-Yellow	23rd St 15-20 mph Delay 3-5 min	A. 3 to 5 B. 15 to 20	A. Heavy B. A short
+6 min	Freeway-Yellow Ramp-Red	23rd St 25-30 mph Delay 5-10 min	A. 5 to 10 B. 25 to 30	A. Moderate B. A 1/2 block
+8 min	Freeway-Red Ramp-Red	23rd St 15-20 mph Delay 5-10 min	A. 5 to 10 B. 15 to 20	A. Heavy B. A 1/2 block

[a]Indications for 27th St. remained yellow.

[b]23rd St. information preceded by

Downtown
X-way
West

and followed by

27th St
25-30 mph
Delay 3-5 min

[c]Roadside radio (tape recording): "At the 23rd Street entrance ramp to the westbound Downtown Expressway, there is a (A) minute delay and the expressway traffic is moving at (B) miles an hour. The 27th Street entrance ramp has a 3 to 5 minute delay, and the expressway speeds are 45 to 55 miles an hour and there is no ramp delay."

[d]Commercial radio (tape recording): "And now, on the westbound Downtown Expressway, traffic is (A) to 25th Street. From there it is moderate to 29th where traffic opens up. There is (B) backup at the 23rd Street entrance and a short backup at the 27th Street entrance ramp."

level of diversion than the others. The chi-square statistic was used to test between pairs of devices at all levels of congestion. Device 1, which shows a sharp dip in the diversion pattern at the 3-min level of delay, proved to be significantly different from all of the other three devices. In fact the only non-significant difference that was noted was between the changeable message matrix sign and the commercial radio. It should be noted that the overall highest level of response was given by the roadside radio. The next highest level of response was given by the changeable message matrix. The abnormal dip in the diversion pattern for the symbolic map may be indicative of the difficulty that a motorist may have in transforming color-coded information into qualitative information for use in his route selection analysis. Although there were significant differences between each of the devices in terms of the diversion pattern, the absolute amount of change in most cases did not exceed 15 percent, indicating that there is little practical difference between the devices. One must then determine whether the device selected to present the information meets an adequate level of reaction and can satisfy other criteria, among them being cost-effectiveness, ease of installation, and maintenance.

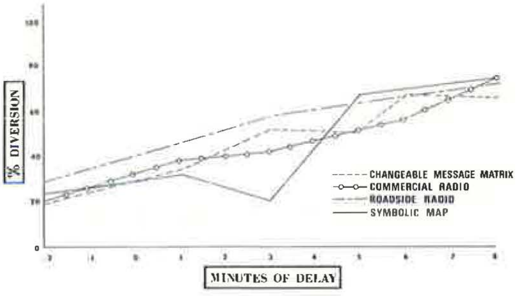


Figure 8. Diversion patterns.

A test of this type gives no real indication of what the motorist would do in terms of actual response to a device in a real driving situation. What has been attempted here is to test reactions under a consistent driving situation to determine whether one device among those tested is either an exceptionally good one, or a very bad one.

From the results obtained it is obvious that the devices performed rather similarly, in that the amount of diversion did tend to increase as the delay caused by using the nearest entrance ramp increased. In most cases it was apparent that the delay or the amount of diversion could possibly be approximated by a linear function of minutes of delay, with the exception of the delay pattern caused by the symbolic map. The symbolic map for that reason may prove to be a somewhat less desirable device than the other three. However, among the other three, it is very difficult to choose one solely on the basis of these data.

Correctness of Response

With data collected under a laboratory situation such as this it is difficult to make rigorous statements concerning the correctness of the response of the individuals to the information. In general, as the amount of delay approached the zero minute level, the diversion rate began to approach 30 to 35 percent. This would indicate that approximately one-third of the population interviewed feels that with no time advantage coming from using the expressway the comfort and convenience offered by the alternate route would be preferable. At the high level of congestion, 8 min of delay, diversion approached 70 percent. Since the shortest alternate route trip was only 16 min, this was a 50 percent increase over the shortest alternate route trip. The diversion for this amount of delay seems reasonable. The problem of the response of the motorists to a symbolic map in the 3-min level of congestion casts some doubt on the ability of this particular device to transmit the proper information.

To summarize, in testing the four devices and the six levels of congestion, it was very difficult to pick out any of the four devices as being superior on the basis of overall response. Except for the symbolic map, all the devices gave a reasonable set of responses to the various level of congestion, and none of the remaining three could be singled out as causing obvious errors in judgment on the part of the respondent.

CONCLUSIONS

This report has attempted to look at the traffic information techniques that might be used to divert drivers around congested parts of the highway system. Because the diversion could both increase the satisfaction of the individual and improve the efficiency of the roadway system, it is important that the most effective technique be utilized. From the survey, it was determined that:

1. Visual techniques were preferred to vocal techniques. The changeable message matrix sign was the most preferred technique.
2. The difference in response level between the various techniques was small, indicating that perhaps any one of them, except the symbolic map, would produce the same results.

When both results are considered, the desirability of more intensive consideration of the changeable message matrix sign becomes apparent. A pilot installation of sufficient magnitude to cover a wide range of traffic conditions and to provide a sufficient data base for evaluation should be implemented to determine the effect of such a system of signs.

REFERENCES

1. Olson, K.W., Sapp, E.R., and Flaig, T.L. A System For Automatic Vehicle Lateral Guidance. Highway Research Record 275, 1969, pp. 1-11.
2. Heathington, Kenneth W. On the Development of a Freeway Driver Information System. Ph.D. dissertation, Northwestern University, 1969.
3. Covault, D.O., and Bowes, R.W. A Study of the Feasibility of Using Roadside Radio Communications for Traffic Control and Driver Information. Highway Research Record 49, 1963, pp. 89-106.
4. McMillan, Robert K., and Assael, Henry. National Survey of Transportation Attitudes and Behavior—Phase I Summary Report. National Cooperative Highway Research Program Report 49, 1968.
5. Eberhard, John W. Driver Information Requirements and Acceptance Criteria. Highway Research Record 265, 1969, pp. 19-30.
6. Wilshire, Roy L. The Benefits of Computer Traffic Control. Traffic Engineering, April 1969.
7. McDermott, Joseph M. Automatic Evaluation of Urban Freeway Operations. Traffic Engineering, Jan. 1968.
8. Benzinger, R.W., and Bell, E. Experimental Route Guidance Head-Up Display Research. Highway Research Record 265, 1969, pp. 62-70.
9. Hoff, Gerald C. Implementation and Evaluation of Experimental Traffic Information Signs. Report 18, Chicago Area Expressway Surveillance Project, 1968.
10. Covault, D.O., Dervish, T., and Kanen, A.C. A Study of the Feasibility of Using Roadside Communications for Traffic Control and Driver Information: Report No. 2. Highway Research Record 202, 1967, pp. 32-66.
11. Wattleworth, Joseph A., and Wallace, Charles E. Evaluation of the Operational Effects of an "On Freeway" Control System. Texas Transportation Institute, 1967.
12. McDermott, Joseph M. General Trends in Peak Period Levels of Expressway Traffic Service. Unpublished draft report, Chicago Area Expressway Surveillance Project, 1969.
13. Thurstone, Leon L. A Law of Comparative Judgment. Psychological Review, Vol. 34, 1927.
14. Torgerson, Warren S. Theory and Methods of Scaling. J.W. Wiley, 1958.
15. Thurstone, Leon L. Rank Order as a Psychophysical Method. Jour. of Experimental Psychology, Vol. 14, 1931.

AN EVALUATION OF A REAR-MOUNTED VEHICLE SPEED INDICATOR

Clinton Jolliffe, Carl Graf, and David Alden,
Honeywell, Inc.

ABRIDGMENT

• THE primary purpose of vehicular taillights at night is to alert drivers to the presence of a vehicle ahead. Once the taillights are detected, the driver processes such visual clues as taillight size, height, motion, luminance level, and apparent inter-taillight separation. Intervehicle distance and speed information is extracted from these clues, and the driver then reacts accordingly.

The existing rear vehicular lighting systems have many shortcomings, as shown by highway accident statistics indicating that 15 percent of all vehicular collisions are of the rear-end type. It has been suggested that a rear signal light should be of dual purpose, i.e., convey vehicle speed information directly in addition to its primary task of alerting the following driver. Hence, it was decided to construct a speed indicator signal system and compare it with a conventional taillight system in terms of a subject's ability to assess the speed and intervehicular distance of a car being overtaken at night.

SPEED INDICATOR DESIGN

The bar-type speed indicator was constructed by mounting 14 amber signal lamps on a board 70 in. long. Each signal lamp had a luminous intensity rating of 2 candlepower. The speed indicator was mounted on the rear of a pickup truck at a height of 45 in. above ground level and 12 in. above the taillights.

The outer two lamps served as reference lamps and were always on. As the speed of the vehicle increased, the signal lights were switched on in pairs at 10-mph increments. That is, for vehicle operation in the 0- to 10-mph speed range, only the two central lamps were on; for the 10- to 20-mph speed range, the four central lamps were on; and so forth until all 12 speed indicator lamps were switched on, which indicated that the vehicle's speed was exceeding 50 mph (Fig. 1). Thus, the width of the bar indicated vehicle speed, and the rate of expansion or contraction corresponded to rate of acceleration or deceleration.

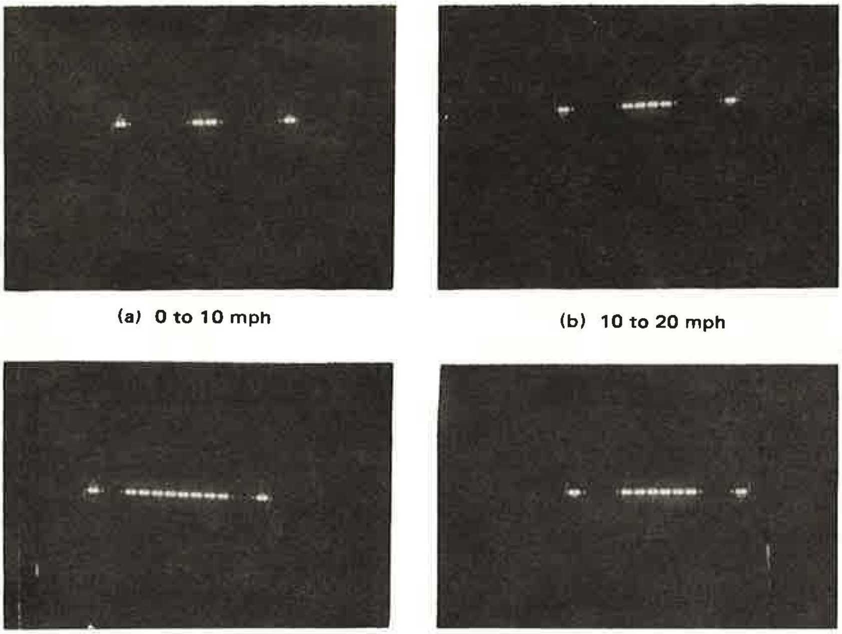
EXPERIMENTAL PROCEDURE

To conduct the study, three vehicles and five persons were employed. The lead vehicle was occupied by a driver and one operator who controlled the signal lights mounted on the rear of the vehicle. Next in line was a "blind" car that had its rear window draped with black cloth. The purpose of this vehicle was to prevent the exposure of the speed indicator or taillight signals to the following subject car until the desired intervehicle distance was achieved. The subject car was occupied by the subject who drove and an experimenter who instructed the subject and recorded the subject's judgments.

The pilot study was conducted at a dragstrip located in a rather isolated rural area; thus there were few extraneous light sources. The dragstrip was 1 mile long and had distance markers posted every 100 ft.

The experiment was structured to obtain subject judgments of the lead car's speed and distance over intervehicle distances that ranged between 100 and 1,000 ft. Four

Sponsored by Committee on Motorist Information Systems.



(a) 0 to 10 mph

(b) 10 to 20 mph

(c) 20 to 30 mph

(d) 30 to 40 mph

Figure 1. Photographs of speed indicator as observed from subject vehicle for four speed conditions.

test speeds (10, 20, 30, and 40 mph) were selected, and the subject was informed that his speed judgment must be one of those assigned speeds. The speed of the subject car was 40 mph under all conditions. Three intervehicle distance categories (far, mid, and near) were established to vary as well as was possible the distances over the specified range. Thus, 24 combinations existed (four speeds, three distance categories, and a taillight or speed indicator). Each condition was run twice for a total of 48 runs per session. The order of the combinations was randomized, with each subject receiving the same random order of runs.

For the test situation, the lead car assumed one of the four selected car speeds. At the appropriate intervehicle distance, the operator signaled the blind car to move into the left lane by turning on either the taillights or speed indicator. The driver in the subject car was then exposed to the rear signal lights for about 2 sec, at which time the rear signal lights were switched off. This signaled the subject to make a speed judgment followed by an intervehicle distance judgment. In addition, this served as a signal for the experimenters in the lead and subject cars to drop a small sandbag from the vehicle's right window for intervehicle distance markers.

Three subjects, ranging in age from 25 to 35 years, were employed for this study. Each subject demonstrated normal color vision, had a visual acuity of 1.0, and had between 10 and 15 years of driving experience.

RESULTS AND DISCUSSION

Chi-square tests indicated no statistically significant relationships between the number of speed judgment errors and intervehicle distance, direction of judgment error and vehicle speed, or distance judgments versus signal type. However, a highly significant relationship ($P < 0.001$) was found between direction of speed judgment error and signal type (Fig. 2). The high significance level was primarily a result of the judgments based on the speed indicator, which were almost consistently overestimated speed judgments.

A sign test of correlated samples indicated that there were significantly fewer speed judgment errors (without regard to sign) with the speed indicator ($P < 0.05$) versus the conventional taillight system.

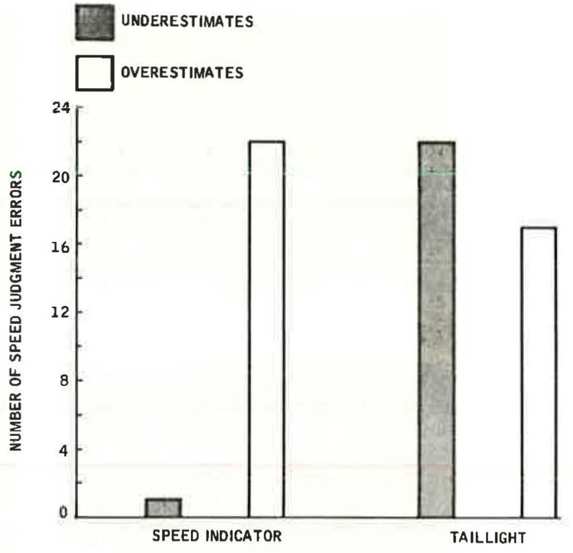


Figure 2. Number and direction of speed judgment errors for each signal type.

One can only speculate as to why subject errors were almost consistently in the direction of overestimation with the speed indicator. Because the subjects were given only a brief orientation period during which time they were able to relate bar length with announced vehicle speed, learning may be a factor. Yet, if it were only a learning factor, one may expect as many errors of underestimation as overestimation. What could have contributed to the bias is that the subjects may also have used apparent vehicle motion as a visual clue to vehicle speed. If this were the case, the speed indicator, due to its greater illumination level, would provide more motion clues for the subject driver and thus account for the evident judgment bias.

As a final comment, we recognize that this pilot study only grazes the surface of the much larger and more complex problem of adequate rear vehicular lighting. We do feel that the concept of a dual-purpose rear signal system of this type is sound and that our findings suggest the need for further study and evaluation.