TRANSPORTATION RESEARCH RECORD 956

Traffic Control Devices and Grade Crossings

TRB
TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL
WASHINGTON, D.C. 1984

Transportation Research Record 956
Price $6.00
Editor: Elizabeth W. Kaplan
Typist: Lucinda Reeder

modes
1 highway transportation
3 rail transportation

subject areas
51 transportation safety
54 operations and traffic control

Transportation Research Board publications are available by ordering directly from TRB. They may also be obtained on a regular basis through organizational or individual affiliation with TRB; affiliates or library subscribers are eligible for substantial discounts. For further information, write to the Transportation Research Board, National Research Council, 2101 Constitution Avenue, N.W., Washington, D.C. 20418

Printed in the United States of America

Library of Congress Cataloging in Publication Data
National Research Council. Transportation Research Board.
Traffic control devices and grade crossings.

(Transportation research record; 956)
1. Traffic signs and signals--Addresses, essays, lectures.
2. Roads--Interchanges and intersections--Addresses, essays, lectures. I. National Research Council (U.S.) Transportation Research Board. II. Title. III. Series.
TE7.H5 no. 956 380.5 s 84-20757
[TE228] [625.7'94]
ISBN 0-309-03701-8 ISSN 0361-1981

Sponsorship of Transportation Research Record 956

GROUP 3–OPERATION AND MAINTENANCE OF TRANSPORTATION FACILITIES
D.E. Orne, Michigan Department of Transportation, chairman

Committee on Traffic Control Devices
Paul H. Fowler, Automobile Club of Southern California, chairman
Philip O. Russell, Federal Highway Administration, secretary
Robert L. Carstens, Robert Dewar, Robert M. Garrett, C. Arthur Geurts, Robert David Henry, Ken F. Kobetsky, Kay K. Krekorian, C. John MacGowan, Hugh W. McGee, Zoltan A. Nemeth, Peter S. Parsonson, Harold C. Rhudy, David G. Snider, Ronald E. Stemmler, Richard H. Sullivan, James I. Taylor, James A. Thompson, Jonathan Upchurch, Clinton A. Venable, W. Scott Wainwright, Earl C. Williams, Jr.

Committee on Railroad-Highway Grade Crossings
John Bradford Hopkins, Transportation Systems Center, chairman
Otto F. Sonefeld, Atchison, Topeka, and Santa Fee Railway Company, secretary
Charles L. Amos, John E. Baerwald, William D. Berg, Thomas M. Bryant, Janet A. Coleman, Denis J. Dwyer, Bruce F. George, William J. Hedley, Frederick J. Kull, Charles Raymond Lewis II, Richard A. Mather, Ernie Oliphant, Edward E. Plasberg, Suzanne M. Roche, Eugene R. Russell, Sr., C. Shoemaker, Barry M. Sweedler, Robert E. Taggart, Jr., Marlen Vanoverbeke, Earl C. Williams, Jr.

David K. Witheford, Transportation Research Board staff

Sponsorship is indicated by a footnote at the end of each report. The organizational units, officers, and members are as of December 31, 1983.

Contents

Addresses of Authors

Barbaresso, James C., Oakland County Road Commission, 31001 Lahser Road, Birmingham, Mich. 48010
Byrd, Michael N., Kimley-Horn and Associates, Inc., 5800 Corporate Way, West Palm Beach, Fla. 33407
Chang, Myung-Soon, Texas Transportation Institute, Texas A&M University, College Station, Tex. 77843
Fambro, Daniel B., Transportation Center, The University of Tennessee, Knoxville, Tenn. 37996-0700
Garber, Nicholas J., Department of Civil Engineering, University of Virginia, Charlottesville, Va. 22901
Heathington, K.W., The University of Tennessee, Knoxville, Tenn. 37996-0700
Hitz, John S., Research and Special Programs Administration, U.S. Department of Transportation, Transportation Systems Center, Cambridge, Mass. 02142
Lin, Feng-Bor, Department of Civil and Environmental Engineering, Clarkson College, Potsdam, N.Y. 13676
Machemehl, Randy B., Department of Civil Engineering, The University of Texas at Austin, Austin, Tex. 78712
Mechler, Ann M., Department of Civil Engineering, The University of Texas at Austin, Austin, Tex. 78712
Messer, Carroll J., Texas Transportation Institute, Texas A&M University, College Station, Tex. 77843
Rochelle, Robert W., Department of Electrical Engineering, The University of Tennessee, Knoxville, Tenn. 37996-0700
Ryan, Timothy A., Kidde Consultants, Inc., 1020 Cromwell Bridge Road, Baltimore, Md. 21204
Saito, Mitsuru, Department of Civil Engineering, University of Virginia, Charlottesville, Va. 22901
Santiago, Alberto, Turner-Fairbank Highway Research Center, FHWA, 6300 Georgetown Pike, McLean, Va. 22101
Stafford, Donald B., Department of Civil Engineering, Clemson University, Clemson, S.C. 29631

Evaluation of Six Active Warning Devices for Use at Railroad-Highway Grade Crossings

K.W. HEATHINGTON, DANIEL B. FAMBRO, and ROBERT W. ROCHELLE

ABSTRACT

Six new active railroad-highway grade crossing warning devices were evaluated under controlled laboratory testing conditions. The six devices included two alternatives for each of three basic systems--four-quadrant gates (with and without skirts), four-quadrant flashing light signals (with and without strobes), and highway traffic signals (with one and with three white bar strobes). The evaluation involved testing the performance of each of the six devices in a near realworld environment to identify the three most desirable devices for subsequent field testing. Thirty-two test subjects drove an instrumented vehicle repeatedly over a private two-lane highway. On each trip down the roadway, the test driver encountered three full-scale active warning devices, any one of which may or may not have been actuated as the vehicle approached. The experimental design included different actuation distances as well as day and night conditions. In addition to driver behavior data, attitudinal data on the effectiveness of the six devices were obtained from each subject. All six active warning devices tested were perceived to be superior to standard active warning devices currently in use at railroad-highway grade crossings. Generally speaking, alternative B of each system (i.e., with skirts, with overhead stobes, and with three white bar strobes) was more effective. Four-quadrant gates with skirts tended to be a superior system in all categories of analysis. The relative effectiveness of flashing light signals and highway traffic signals tended to alternate depending on the category of analysis; there was not a consistent ordering of effectiveness of these two systems.

Research to improve safety at railroad-highway grade crossings has been going on for some 50 years, but the methods used for warning motorists of impending danger at a crossing have not changed significantly. During this time many innovative warning systems have been developed for use both at and in advance of crossings and millions of dollars have been spent in developing, testing, and evaluating these devices. Yet field implementation of new concepts has been minimal.

As part of a research project that addresses these issues, eight innovative active warning devices were identified as having potential for improving safety at railroad-highway grade crossings (1). A prioritization of these eight devices identified five of them for detailed laboratory evaluation. However, careful review of these five devices determined that they were in fact variations of three conceptually different systems (i.e., gates, flashing light signals, and highway traffic signals). For this reason, it was proposed that laboratory testing evaluate six devices consisting of two variations of each of the three basic systems. The devices chosen for testing were

1. Four-quadrant gate system with and without skirts,
2. Four-quadrant flashing light signal system with and without overhead strobes, and
3. Highway traffic signal system with one and with three white bar strobes.

The evaluation process involved testing the performance of each of these devices in a near real-world environment to identify the three most desirable devices for subsequent field testing. The configuration of each prototype device was in accordance with the Manual on Uniform Traffic Control Devices (2) and standard highway engineering practice. Figures 1-3 show the installation of the six devices as they were evaluated in the laboratory testing. The results of this study are summarized in this paper; supporting documentation is contained in another report (3).

EXPERIMENTAL PLAN

The laboratory evaluation of the active warning devices involved 32 test subjects each of whom drove

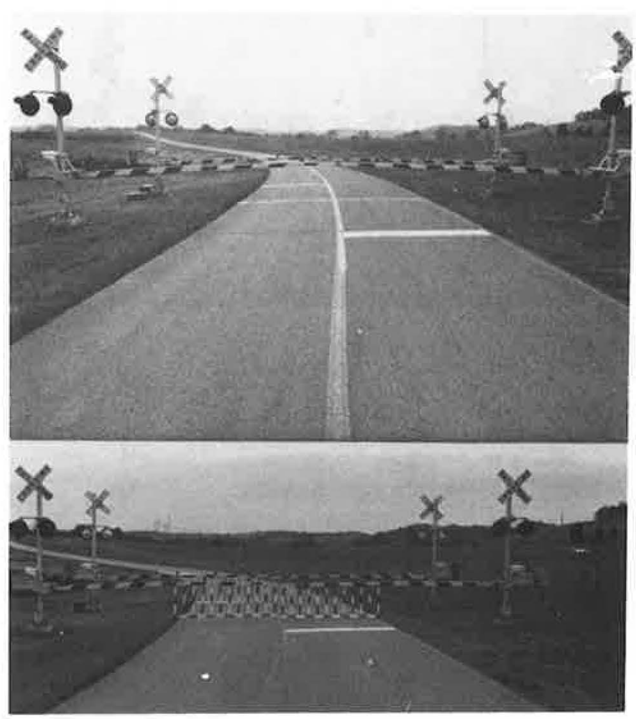

FIGURE 1 Four-quadrant gate system: top, alternative A, without skirts on gate arms; bottom, alternative B, with skirts added to all gate arms.

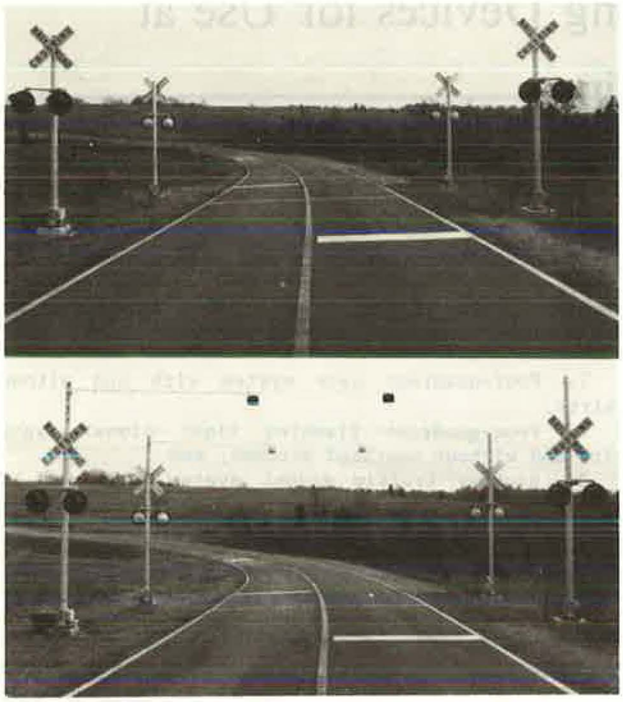

FIGURE 2 Four-quadrant flashing light signal system: top, alternative A, without strobes; bottom, alternative B, with red strobe lights over traffic lanes.

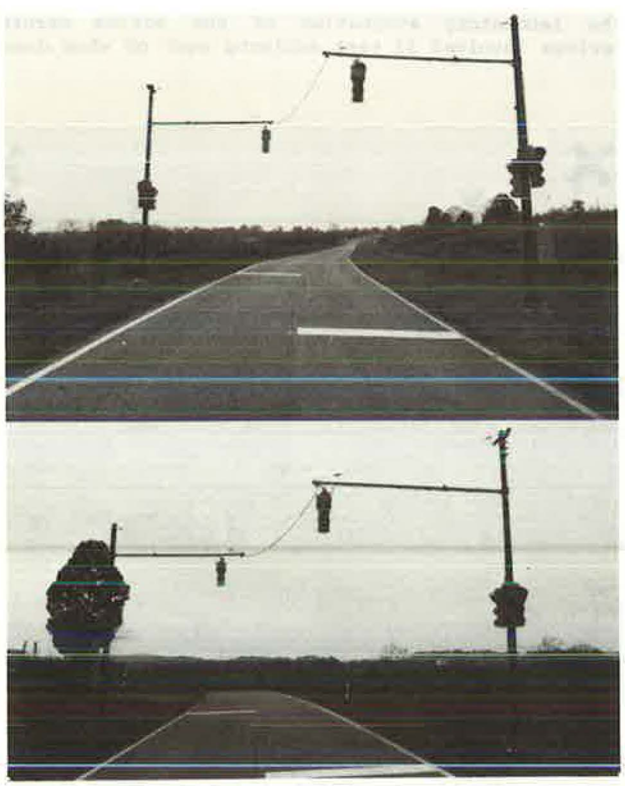

FIGURE 3 Highway traffic signal system: top, alternative A, white bar strobe in overhead red signal lens; bottom, white bar strobe in each red signal lens.

an instrumented vehicle at approximately 40 mph repeatedly over a 1.5-mile stretch of private two-lane highway. On each trip down the roadway, the test driver encountered three full-scale active railroad-highway grade crossing warning devices, any one of which may or may not have been actuated as the vehicle approached. Subjects were requested to respond to the traffic control devices as they would under normal driving conditions. An experimental design was developed in which the effects of the following independent variables could be evaluated (4):

1. Alternative active warning devices (alternative A versus alternative B for each system);
2. Basic active warning systems (system A versus system B versus system C);
3. Signal actuation distance (null, long, medium, and short); and
4. Lighting conditions (day versus night).

Each subject experienced two replications of each of the 48 treatment combinations. All of the variables are self-explanatory with the exception of actuation distance that was defined as the distance the test vehicle was from the device when the device communicated a changing of right-of-way (i.e., flashing light signals were actuated or the highway traffic signal changed to yellow).

RESULTS

Subject Characteristics

Half of the 32 subjects were male and half were female; each group was further divided into an equal number of younger (under 25) and older (over 60) drivers. The younger subjects had been driving an average of 4.1 years and the older drivers an average of 45.8 years. In both age groups, the males drove almost twice as many miles per year as did their female counterparts. The average educational level of the subject population was similar for each of the four groups; however, individuals within the groups ranged from those who did not complete high school to college graduates.

When the subjects' simple reaction times were measured, the young male subjects were the fastest (0.41 sec) of any of the four groups. Times for the other three categories were all about 0.48 sec. When vision was tested, the average corrected visual acuity of the younger subjects was about 20/20. For the older subjects, the average was about 20/30, which is the legal requirement for operating a motor vehicle in Tennessee. However, two younger and five older drivers did not meet this criterion.

Attitudinal Data

After completing all of the test runs, each subject was asked to compare the effectiveness of each of the six alternative devices with that of existing signals. An absolute scale of 1 to 5 was used for this purpose, with 3 being about the same as existing signals, 1 being much less effective, and 5 being much more effective. Responses were solicited for both day and night driving conditions. All six of the devices were ranked higher than existing signals in both situations. In each case, the four-quadrant flashing lights was the lowest ranked alternative. The rank order of the other four devices varied by time of day. For day driving, the order was four-quadrant gates, highway traffic signals,

and four-quadrant flashing light signals. However, this was not true for night; one of the four-quadrant flashing light signal alternatives was rated as the second most effective device at night.

In addition to the absolute rankings, the subjects were asked to pick the more effective device from a series of two alternatives. Thurstone's method of paired comparisons (5) was used to determine a relative ranking of the six alternative devices. Figure 4 shows the results of these rankings for both day and night conditions. As shown, the two highest ranked alternatives involved four-quadrant gates with and without skirts. The third ranked device was the four-quadrant flashing light signals with overhead strobes. The next two ranked devices involved the highway traffic signal. The lowest ranked alternative was the four-quadrant flashing light signals without overhead strobes. This ranking technique clearly shows the gate system with skirts to be much preferred to the other systems.

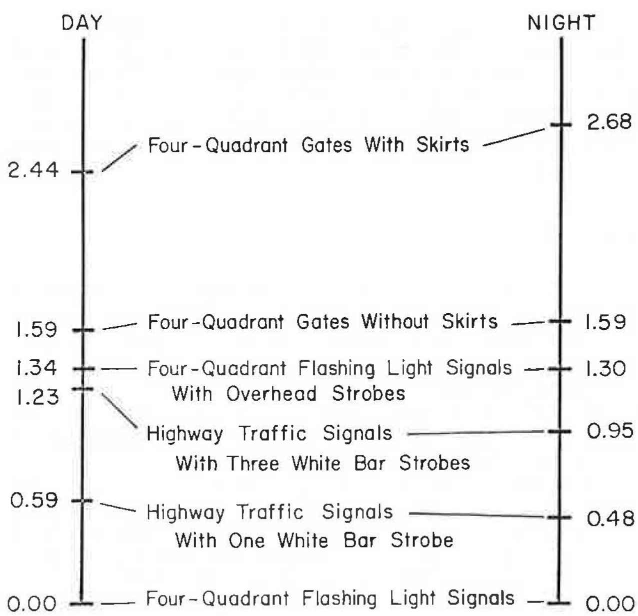

FIGURE 4 Relative ranking of the six alternative active warning devices.

Driver Behavior Data

To investigate the effects of the six innovative devices on subject behavior, each of 3,024 observations was classified according to device type, actuation distance, and lighting condition. This resulted in a model with 48 cells (6 devices x 4 distances x 2 conditions). Brake reaction time and maximum deceleration rate were selected as the most meaningful variables from the candidate list of variables describing a subject's reaction to a traffic control device. The cell means for these variables were calculated and stored in tabular format. An attempt was made to analyze the number of extreme deceleration rates in the data set; however, because only 13 of the observed rates could be classified as extreme, a further analysis was not conducted.

To determine whether the cell mean differences were real or simply a result of chance, three-way analysis of variance (ANOVA) tests for statistical significance were performed. Duncan's test (6) was used to test the differences for each of the three main effects in the full model. Results are as follows:

1. When response to the six device types is averaged across all levels of actuation distance and lighting condition, there is not a significant difference between the two gate systems and the two flashing light systems; however, the difference in response to these four devices and to the two highway traffic signal systems is significant.
2. When response to four actuation distances is averaged across all levels of device type and lighting condition, there are significant differences between each of the four distances.
3. When response to the two lighting conditions is averaged across all levels of device type and actuation distance, there is significant difference between day and night; however, the magnitude of the difference (0.23 sec) is very small.

Because actuation distance explained most of the variation in the data set, it was decided to run separate two-way ANOVA tests for each of the four actuation distances. The results are summarized in Table 1. Interpretation of this table is rather complex because both variables are represented and each column depicts a separate ANOVA test. There are only 12 cells (6 devices x 2 conditions) in each model, and there are 4 models for each variable. For each model, cells with the same letter are not significantly different.

As an example, for the brake reaction time variable and the medium actuation distance, response to the two types of four-quadrant gates is significantly different (faster) than response to the other four devices; however, differences within the two groups (two types of gates and the other four devices) are not statistically significant. For the short actuation distance, response to the four-quadrant gates with skirts is significantly different from response to the two types of highway traffic signals. In addition, response to the two types of four-quadrant gate systems and the two types of four-quadrant flashing light signal systems is significantly different from the response to two types of highway signals. Because the four-quadrant gates without skirts belong to both groups A and B, there is no significant difference between the four-quadrant gate systems and four-quadrant flashing light signal systems.

Basically, the four-quadrant gates with skirts always belong to the fastest group, and the two types of highway traffic signals are always associated with the slowest. The placement of the four-quadrant gates without skirts and the two types of four-quadrant flashing light signals varies with actuation distance. In addition, at both the medium and long distances, response to the devices is significantly different between day and night conditions.

Interpretation of the maximum deceleration rates is not as clear-cut. There is not a significant difference in deceleration at the short actuation distance, but at the other distances there are differences. However, there appears to be no pattern to the results. As with the brake reaction time variable, response to the devices at the medium and long actuation distances was significantly different under day and night conditions.

CONCLUSIONS

All six innovative active warning devices were perceived as superior to standard active warning devices currently in use at railroad-highway grade crossings. The subjects always perceived the four-quadrant gates with skirts as the most effective and four-quadrant flashing light signals without strobes as the least effective on both an absolute and rela-

TABLE 1 Comparison of Average Brake Reaction Times and Maximum Deceleration Rates for Various Actuation Distances

		Actuation Distance			
Variable	Type of Traffic Control	Short (330 ft)	Medium[a] (440 ft)	Long[a] (670 ft)	Null (0 ft)
First brake, time	Four-quadrant gates without skirts	A,B	A	A	A
	Four-quadrant flashing light signals without strobes	B	B	A	A
	Highway traffic signals–one white bar strobe	C	B	B	B
	Four-quadrant gates with skirts	A	A	A	A
	Four-quadrant flashing light signals with overhead strobes	B	B	A	A
	Highway traffic signals–three white bar strobes	C	B	B	B
Maximum deceleration, rate	Four-quadrant gates without skirts	A	A,B	A	A,B
	Four-quadrant flashing light signals without strobes	A	C	B	B,C
	Highway traffic signals–one white bar strobe	A	A,B,C	A	A
	Four-quadrant gates with skirts	A	B,C	A	A
	Four-quadrant flashing light signals with overhead strobes	A	B,C	B	C
	Highway traffic signals–three white bar strobes	A	A	A	A

Note: Cells with the same letter are not significantly different.
[a]Significant differences between day and night.

tive ranking basis. The perceived effectiveness of the other four devices tended to cluster together and was not significantly different on the absolute ranking basis for either day or night conditions. However, on the relative ranking basis, there was a consistent order of perceived effectiveness for both conditions.

At both the short and medium actuation distances, four-quadrant gates with skirts resulted in significantly quicker brake reaction times than either highway traffic signals or four-quadrant flashing light signals. In addition, at all actuation distances, the two highway traffic signal alternatives resulted in significantly slower brake reaction times than either of the other two systems. At the short actuation distance, there was no significant difference in the resultant deceleration rates for any of the six active warning devices, but there was for medium and long activation distances.

Generally speaking, alternative B of each system (i.e., with skirts, with overhead strobes, and with three white bar strobes) was more effective; four-quadrant gates with skirts tended to be a superior system in all categories of analysis; and the effectiveness of four-quadrant flashing light signals and highway traffic signals tended to alternate relative to one another depending on the category of analysis; there was not a consistent ordering of the effectiveness of these two systems.

RECOMMENDATIONS

Because the overall research project is directed toward determining the cost-effectiveness of alternative active warning devices and the cost of each basic system tested varies over a wide range, it is recommended that one alternative of each basic system be field tested. Because alternative B of each basic system was generally superior, it should be the one that is field tested. One field test of each alternative should be begun initially rather than two installations of each alternative. Field experience is needed before additional crossings are implemented.

ACKNOWLEDGMENTS

The authors wish to acknowledge the contributions of Janet Coleman of the FHWA. Special thanks are given to the U.S. Air Force for the use of the roadway facilities at the McGhee-Tyson Air National Guard Base. In addition, the authors are greatly indebted to the 32 test subjects, who often spent much more time at the test site than anticipated. Without their cooperation, this study would not have been successful.

REFERENCES

1. J.C. Hungerford and K.W. Heathington. Conceptualization and Subjective Evaluation of Innovative Railroad-Highway Crossing Active Warning Devices. Transportation Center, University of Tennessee, Knoxville, 1981.
2. Manual on Uniform Traffic Control Devices for Streets and Highways. FHWA, U.S. Department of Transportation, 1978.
3. K.W. Heathington, D.B. Fambro, and R.W. Rochelle. Laboratory Evaluation of Six Active Warning Devices for Use at Railroad-Highway Grade Crossings. Transportation Center, University of Tennessee, Knoxville, July 1983.
4. J.N. Snider and J.C. Hungerford. Laboratory Specification and Testing of Alternative Active Railroad-Highway Grade Crossing Warning Devices. Transportation Center, University of Tennessee, Knoxville, Jan. 1982.
5. L.L. Thurstone. The Measurement of Values. The University of Chicago Press, Chicago, 1959.
6. D.B. Duncan. t-Tests and Intervals for Comparisons Suggested by the Data. Biometrics, Vol. 31, June 1975, pp. 339-359.

This paper is the result of work of researchers at The University of Tennessee who are involved in FHWA contract DTFH61-80-C-00180, Innovative Railroad-Highway Crossing Active Warning Devices. The contents of this paper reflect the views of the authors who are responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the FHWA.

Publication of this paper sponsored by Committee on Railroad-Highway Grade Crossings.

Accident Severity Prediction Formula for Rail-Highway Crossings

JOHN S. HITZ

ABSTRACT

The development of formulas to predict the severity of accidents at public rail-highway crossings is described. The formulas make use of the previously developed DOT Accident Prediction Formula, the U.S. DOT-AAR National Rail-Highway Crossing Inventory, and the FRA accident files. When these new formulas are used in the DOT Resource Allocation Procedure, information will be available to assist in making better decisions about where to install motorist-warning devices to further increase crossing safety for a given level of funding. Established statistical techniques were used to develop two formulas: one that estimates the number of fatal accidents per year at a crossing and one that estimates the number of injury accidents per year at a crossing. It was found that the factors in the inventory that significantly influence fatal accident severity, given that an accident occurred, were maximum timetable train speed, the number of through trains per day, the number of switch trains per day, and urban or rural location. For injury accident severity, given that an accident occurred, the significant factors were maximum timetable train speed, the number of tracks, and urban or rural location. The performance of these severity formulas is discussed and calculated results are presented.

The DOT Rail-Highway Crossing Resource Allocation Procedure, developed at the U.S. Department of Transportation's Transportation Systems Center (TSC), employs an accident prediction formula. In an attempt to improve the effectiveness and usefulness of the resource allocation procedure, a study was undertaken to incorporate a quantitative measure of severity in the accident prediction formula. That study (1) is documented in this paper. Two severity formulas were developed using established statistical techniques; one formula estimates the number of fatal accidents per year at a crossing and the other estimates the number of injury accidents per year at a crossing. The resulting formulas are to be incorporated in the DOT Rail-Highway Crossing Resource Allocation Procedure (2,3).

BACKGROUND

The Highway Safety Acts of 1973 and 1976 and the Surface Transportation Assistance Acts of 1978 and 1982 provide federal funding authorizations to states specifically for safety improvement projects at public rail-highway crossings. Such safety improvements frequently involve the installation of active motorist-warning devices such as flashing lights or gates. To promote the effective use of federal funds for these safety projects, the U.S. Department of Transportation (DOT) has developed a procedure to assist states and railroads in planning rail-highway crossing safety programs. This procedure, the DOT Rail-Highway Crossing Resource Allocation Procedure (DOT procedure), determines crossing safety improvements that result in the greatest accident reduction benefits based on consideration of predicted accidents at crossings, the costs and effectiveness of safety improvement options, and budget limits.

Two analytic methods have been developed as part of the DOT procedure. Their development followed completion of a joint U.S. DOT-Association of American Railroads (AAR) National Rail-Highway Crossing Inventory (inventory), which numbered and collected inventory information for all public and private crossings in the United States (4). The first analytic method included in the DOT procedure is the DOT accident prediction formula, which computes the expected number of accidents at crossings based on information available in the inventory and crossing accident data files (5). The second analytic method is a resource allocation model designed to rank crossings that are candidates for improvement on a cost-effective basis and to recommend the type of warning device that is to be installed (6).

The current effort is motivated by the recognition that not all rail-highway crossing accidents are equally severe. In 1981 there were a total of 8,546 rail-highway crossing accidents (7). Of these accidents 5,761 caused no casualties, 2,224 caused injuries only, and 561 involved fatalities. Thus, 67 percent of the accidents involved no measurable casualty severity, and only 6.6 percent involved fatalities. This unequal distribution of severity among crossing accidents makes it important, but difficult, to identify those crossings that are likely to have high-severity accidents. A priority ranking of crossings by number of predicted accidents (as done by current DOT procedure) could be significantly different from such a ranking by predicted severity of accident. This difference might affect the use of improvement funds.

ACCIDENT SEVERITY FORMULA

The traditional approach to risk analysis (8) views safety risk as the product of two independent factors: (a) the frequency of accident occurrence, and (b) the severity or consequences of accident occurrence. The product of these two factors for a given hazardous situation provides the total safety risk for that situation. For example, a rail-highway crossing with a predicted accident frequency of 0.5 accidents per year and a predicted accident severity of 0.2 fatalities per accident poses a total safety risk of 0.1 fatalities per year. The division of safety risk into accident frequency and severity

components is particularly appropriate for the current effort because one of the components, the DOT accident prediction formula, already exists. The proposed severity prediction formula would be used with the accident prediction formula to provide a prediction of total safety risk as follows:

$$R = A \times S \quad (1)$$

where

- R = risk of a crossing measured in expected casualties per year,
- A = predicted accident frequency from the current DOT accident prediction formula, and
- S = predicted accident casualties from the severity prediction model.

A major benefit of this approach is that the current DOT accident prediction formula will remain unchanged and can be used either with or without the severity formula. Procedures for use of the severity formula with the DOT accident prediction formula and the DOT Rail-Highway Crossing Resource Allocation Procedure will be described in an updated version of the Rail-Highway Crossing Resource Allocation Procedures User's Guide (3), due for completion during fiscal year 1984.

APPROACH

Under this effort two severity formulas were developed: one formula to predict fatality severity and another to predict injury severity. These formulas provide predictions on the basis of the crossing characteristics described in the inventory. The first task in developing the severity formulas involved the selection of specific measures of severity to be quantified by the formulas. The next task was to identify in the inventory crossing factors that showed a strong correlation with measures of severity for possible inclusion in the severity formulas. The severity formulas were then developed using a regression procedure, referred to as the logistic discriminant approach, which employs an iterative weighted regression technique that is a modification of a method described in Cox (9). The last task in development of the severity formulas was to evaluate their performance by comparing predicted versus actual accident severity.

SEVERITY PREDICTION FORMULA DEVELOPMENT

Selection of Severity Measures

The proposed use of the severity formulas dictates that severity be measured in terms of consequences, given that an accident occurred. The severity measures must therefore be expressed in terms of consequences per accident. The current effort concentrated on developing formulas for quantifying fatalities and injuries as measures of severity.

For the purposes of this study, a fatal accident is an accident in which at least one fatality occurred independent of injuries or property damage; an injury accident is an accident in which there were no fatalities and at least one injury occurred independent of property damage.

To assist in evaluating alternative measures of fatality and injury severity, histograms were developed as shown in Figures 1 and 2. These histograms relate average values of the measures, calculated from accident records, to accidents grouped by intervals of maximum train speed. This permits a review of how the measures vary as a function of a factor (maximum timetable train speed) previously shown to be correlated with accident severity (10). It should be noted that maximum timetable train speed is a crossing characteristic included in the inventory, and it is used here as a surrogate for actual train speed at the time of an accident.

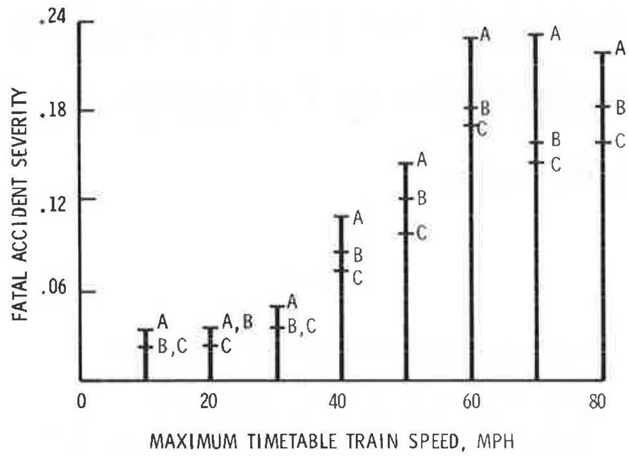

FIGURE 1 Comparison of fatality severity measures.

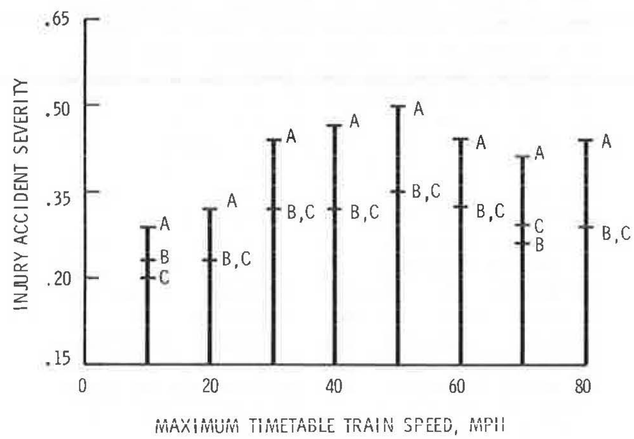

FIGURE 2 Comparison of injury severity measures.

The histograms in Figure 1 show that the three fatality measures considered vary with train speed in the same general manner. All three increase with train speed to about 60 mph beyond which they remain relatively constant. This is intuitive because, beyond some high value of severity, fatalities can no longer increase. As originally surmised, values for fatalities per accident are higher than values for fatal accidents per accident which, in turn, are higher than those for fatalities per occupant per accident. The shape of the histograms for the three measures is generally the same, however, suggesting that any of the measures could be used with similar results. Given the general compatibility of the measures, fatal accidents per accident was chosen as the measure of fatality severity because it avoids the complexities of dealing with vehicle occupants.

This measure can be restated, in statistical terms, as the probability of a fatal accident, given an accident.

The same reasoning led to the selection of injury accidents per accident as the measure of injury severity. This measure can be restated as the probability of an injury accident, given an accident. It is of interest to note from Figure 2 that the shape of the injury severity histograms increases and then decreases with increasing train speed. This is also intuitive because, beyond some severity threshold, casualties will increasingly be fatalities rather than injuries.

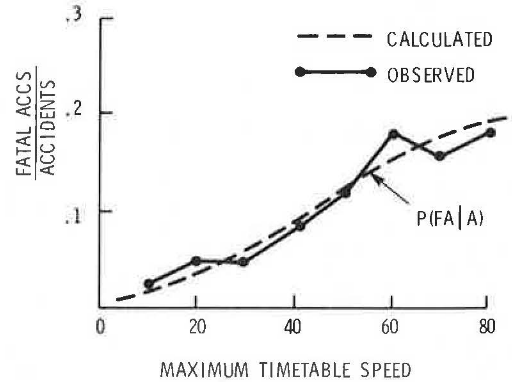

FIGURE 3 Typical plot of observed fatal accident frequency and calculated values of P(FA|A).

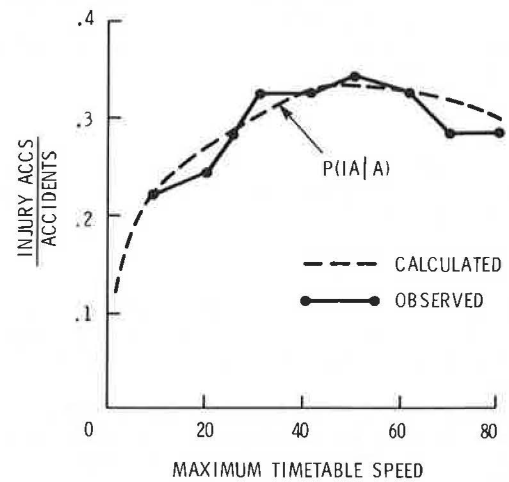

FIGURE 4 Typical plot of observed injury accident frequency and calculated values of P(IA|A).

Selection of Severity Factors

Development of the severity formulas started with identification of factors that correlate with the severity measures and are thus potential predictors of severity. All crossing characteristic factors in the inventory were systematically reviewed to identify those correlated with the severity measures. To accomplish this, histograms similar to Figures 1 and 2 were developed relating average values of the measures calculated for accidents grouped by intervals of the factor in question. Results of this analysis showed that train speed was the strongest predictor of fatal accident severity of all the factors in the inventory. This is consistent with results obtained by Coleman and Stewart (10) in an earlier study of crossing accident data. Histograms were also constructed relating the severity measures to two factors. The following factors were identified as potentially useful in predicting fatality and injury severity:

-Maximum timetable train speed,
-Urban or rural crossing,
-Number of main tracks,
-Number of other tracks,
-Number of through trains, and
-Number of switch trains.

Summary of Formula Development

The analytic objective of this phase of the study was to develop formulas that would predict the probability of a fatal accident given an accident, P(FA|A), and the probability of an injury accident given an accident, P(IA|A). From these two formulas the safety risk expressed in terms of expected number of fatal accidents, R_f, and injury accidents, R_i, per year at a crossing can be determined from

$$R_f = A \times P(FA|A) \quad (2)$$

$$R_i = A \times P(IA|A) \quad (3)$$

where A is the expected number of accidents per year at the crossing from the DOT accident prediction formula.

The analytic character of the fatal accident probability function, P(FA|A), relative to observed data is shown in Figure 3. This graph is a frequency plot of the observed ratio of fatal accidents to total accidents versus maximum timetable train speed. The function P(FA|A) is represented by the dashed line that is a best fit to the observed data points connected by the solid line. Of course, the severity formula is multivariate and, hence, the dashed line for P(FA|A) would be a multidimensional "surface."

The analytic character of P(IA|A) relative to observed data is shown in Figure 4. This graph is a frequency plot of the observed ratio of injury accidents to total accidents versus the same variable, maximum timetable train speed. In this case, the function P(IA|A) does not increase monotonically with severity. However, the particular regression procedure used to develop the severity formulas involved fitting a monotonic function to the observed data. The required formula for predicting injury accident probability could, therefore, not be obtained directly from the regression analysis. This problem was overcome by limiting the accident data to nonfatal accidents. A formula was then developed, from the regression analyses, to predict the probability of an injury accident given that a nonfatal accident occurred, P(IA|NFA). The formula for P(IA|NFA) is, as required, a monotonically increasing function of severity. Having obtained the formula for P(IA|NFA), the desired formula for P(IA|A) was then obtained from the following relationship:

$$P(IA|A) = P(IA|NFA) \times P(NFA|A) \quad (4)$$

where P(NFA|A) is the probability of a nonfatal accident, given an accident, that is,

$$P(NFA|A) = 1 - P(FA|A) \quad (5)$$

where P(FA|A) is the fatal accident probability formula obtained earlier. Hence,

$$P(IA|A) = P(IA|NFA) \times [1 - P(FA|A)] \quad (6)$$

In performing the regression analyses, the observed data for the dependent variable were assigned only two values. In the case of the fatal accident formula these values were +1 for a fatal accident and -1 for a nonfatal accident. For the injury accident formula the values assigned were +1 for an injury accident and -1 for a noninjury accident. The data used for the analyses were for the years 1978-1980. The regression analyses resulted in nonlinear formulas for the dependent variable f, from the fatal accident data, and i, from the injury accident data.

The resulting regression formulas typically produced values between +1 and -1 for the independent variables f and i. Extreme values of the independent variables f and i can, in theory, be from +00 to -00. The desired values for f and i, however, are between 0 and 1 as required by the probability functions P(FA|A) and P(IA|A). The formulas for f and i therefore had to be transformed into probability functions. To accomplish this the following transformation was made to f to obtain the desired fatal accident probability formula:

$$P(FA|A) = 1/(1 + e^{-2f}) \quad (7)$$

For the injury accident formula, the probability of an injury accident given a nonfatal accident, P(IA|NFA), was obtained first:

$$P(IA|NFA) = 1/(1 + e^{-2i}) \quad (8)$$

The probability of an injury accident given an accident, P(IA|A), was then obtained by substituting Equations 7 and 8 into Equation 6 as described previously.

This discussion has provided an overview of the strategy involved in obtaining the formulas required for predicting fatal accident and injury accident probabilities. A more detailed discussion of the regression analysis is presented elsewhere (1).

Resulting Severity Prediction Formulas

The resulting formulas for predicting the probabilities of fatal accidents and injury accidents can be expressed in terms of several factors that are combined by simple mathematical operations. Each factor in the formulas represents a crossing characteristic described in the inventory. The probability of a fatal accident given an accident, P(FA|A), is expressed as

$$P(FA|A) = 1/(1 + CF \times MS \times TT \times TS \times UR) \quad (9)$$

where

CF = formula constant = 695,
MS = factor for maximum timetable train speed,
TT = factor for through trains per day,
TS = factor for switch trains per day, and
UR = factor for urban or rural crossing.

The equations for calculating crossing characteristic factors for the fatal accident probability formula are

$$CF = 695$$
$$MS = ms^{-1.074}$$
$$TT = (tt + 1)^{-0.1025}$$
$$TS = (ts + 1)^{0.1025}$$
$$UR = e^{0.188ur}$$

where

ms = maximum timetable train speed (mph);
ts = number of switch trains per day;
tt = number of through trains per day; and
ur = 1 for urban crossing, 0 for rural crossing.

The probability of an injury accident given an accident, P(IA|A), is expressed as

$$P(IA|A) = [1 - P(FA|A)]/(1 + CI \times MS \times TK \times UR) \quad (10)$$

where

P(FA|A) = probability of a fatal accident, given an accident, obtained from Equation 9,
CI = formula constant = 4.280,
MS = factor for maximum timetable train speed,
TK = factor for number of tracks, and
UR = factor for urban or rural crossing.

The equations for calculating crossing characteristic factors for the injury accident probability formula are

$$CI = 4.280$$
$$MS = ms^{-0.2334}$$
$$TK = e^{0.1176tk}$$
$$UR = e^{0.1844ur}$$

where

ms = maximum timetable train speed (mph);
ur = 1 for urban crossing, 0 for rural crossing; and
tk = total number of tracks at crossing.

To simplify use of the formulas, the values of the crossing characteristic factors have been tabulated for typical values of crossing characteristics. These values are given in Tables 1 and 2 for the fatal accident and injury accident probability formulas, respectively.

Use of Severity Prediction Formula

A sample application of the fatal and injury accident severity formula for a typical crossing is provided to demonstrate their use. Characteristics of the sample crossing are listed in Table 3.

To determine the probability of a fatal accident given an accident at the sample crossing, Equation 9 is used. Values for the factors in the fatal accident severity formula (Equation 9) can be computed from the equations given previously or looked up in Table 1. Table 1 gives the following factor values for the crossing characteristics specified:

CF = 695.0
MS = 0.019
TT = 0.782
TS = 1.202
UR = 1.000

Substituting the factor values into the fatal accident probability formula yields

$$\begin{aligned} P(FA|A) &= 1/(1 + CF \times MS \times TT \times TS \times UR) \\ &= 1/(1 + 695.0 \times 0.019 \times 0.782 \times 1.202 \times 1.000) \\ &= .075 \end{aligned}$$

To determine the probability of an injury acci-

TABLE 1 Factor Values for Fatal Accident Probability Formula

Formula Constant (CF)	Maximum Timetable Train Speed (mph)	MS	No. of Through Trains/Day	TT	No. of Switch Trains/Day	TS	Urban or Rural Crossing[a]	UR
695.0	1	1.000	0	1.000	0	1.000	0	1.000
	5	0.178	1	0.931	1	1.074		
	10	0.084	2	0.894	2	1.119	1	1.207
	15	0.055	3	0.868	3	1.152		
	20	0.040	4	0.848	4	1.179		
	25	0.032	5	0.832	5	1.202		
	30	0.026	6	0.819	6	1.221		
	40	0.019	7	0.808	7	1.238		
	50	0.015	9	0.790	9	1.266		
	60	0.012	10	0.782	10	1.279		
	70	0.010	20	0.732	20	1.366		
	80	0.009	30	0.703	30	1.422		
	90	0.008	40	0.683	40	1.464		
	100	0.007	50	0.668	50	1.497		

[a] 0 = rural, 1 = urban.

TABLE 2 Factor Values for Injury Accident Probability Formula

Formula Constant (CI)	Maximum Timetable Train Speed (mph)	MS	Total Number of Tracks	TK	Urban or Rural Crossing[a]	UR
4.280	1	1.000	0	1.000	0	1.000
	5	0.687	1	1.125	1	1.202
	10	0.584	2	1.265		
	15	0.531	3	1.423		
	20	0.497	5	1.800		
	25	0.472	6	2.025		
	30	0.452	7	2.278		
	40	0.423	8	2.562		
	50	0.401	9	2.882		
	60	0.385	10	3.241		
	70	0.371	15	5.836		
	80	0.360	20	10.507		
	90	0.350				
	100	0.341				

[a] 0 = rural, 1 = urban.

TABLE 3 Characteristics of Sample Crossing

Characteristic	Value
Maximum timetable train speed (mph)	40
Through trains per day	10
Switch trains per day	5
Total number of tracks (main plus other)	2
Urban or rural location	Rural

dent given an accident, at the same sample crossing, Equation 10 is used. Values for the factors in Equation 10 can be obtained from the equations given previously or from Table 2. Table 2 gives the following factor values for the characteristics of the sample crossing:

P(FA|A) = .075 (from fatal accident severity formula)
CI = 4.280
MS = 0.423
TK = 1.265
UR = 1.000

Substituting the factor values into the injury accident probability formula yields

$$\begin{aligned} P(IA|A) &= [1 - P(FA|A)]/(1 + CI \times MS \times TK \times UR) \\ &= (1 - .075)/(1 + 4.280 \times 0.423 \times 1.265 \times 1.000) \\ &= 0.281 \end{aligned}$$

SEVERITY FORMULA PERFORMANCE

To illustrate characteristics of the fatal and injury severity formulas, the two functions P(FA|A) and P(IA|A) are plotted as a function of maximum timetable train speed and one other severity factor in Figures 5 and 6. The probability of a fatal accident given an accident P(FA|A) (Figure 5) increases as a nearly linear function of timetable train speed. Changes in the number of through trains do not have a major influence on fatal accident severity. The probability of an injury accident given an accident P(IA|A) (Figure 6) increases as a nonlinear function of timetable train speed. Injury accident severity generally increases rapidly with timetable train speed and then remains relatively constant beyond 40 mph. The function actually decreases at high speeds under certain conditions as previously predicted from observation of actual accident data (see Figure 4). The number of tracks at the crossing has a significant influence on the function (injury accident severity decreases with the number of tracks).

The performance of the severity formulas was evaluated using two methods: (a) comparing predicted versus actual severity for sample sets of accidents and (b) comparing the ability of the formulas to rank accidents by severity with a random ranking. Results of the first evaluation are summarized in Table 4. Using 1978, 1979, and 1980 data, the severity formulas were used to predict the number of

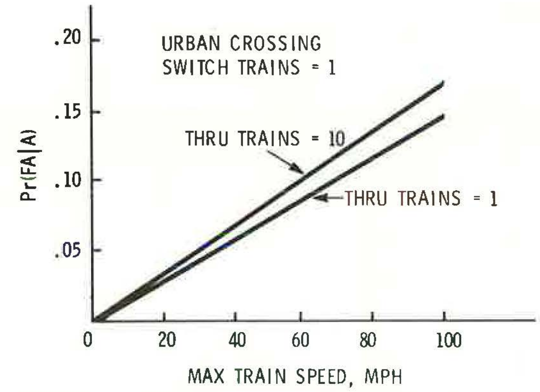

FIGURE 5 Probability of fatal accident, given an accident, versus maximum timetable train speed.

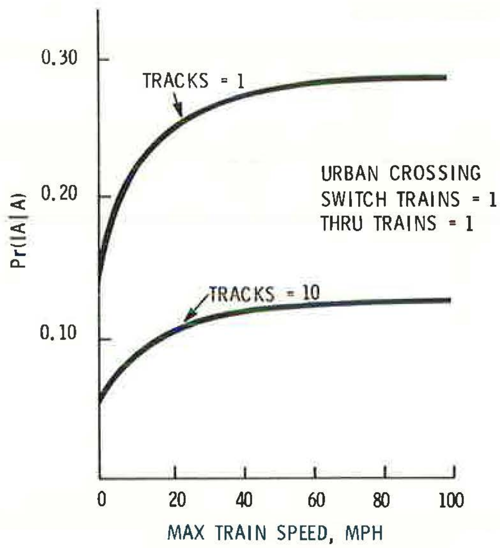

FIGURE 6 Probability of injury accident, given an accident, versus maximum timetable train speed.

TABLE 4 Predicted Versus Actual Accident Severity

No. of Ranked Accidents	No. of Predicted Fatal Accidents	No. of Actual Fatal Accidents	No. of Predicted Injury Accidents	No. of Actual Injury Accidents
100	18.2	13	31.3	42
500	79.3	76	154.2	171
1,000	142.6	145	305.9	348
7,934	511.9	539	2,018.5	2,192

fatal and injury accidents for sets of accidents that occurred in 1981. The predictions were then compared with actual accident records for the same sets of accidents. The sets of accidents considered were selected from the top of a list of accidents ranked by predicted severity. According to Table 4 the severity prediction formulas compare well with observed data. For example, the first row shows that, for the top 100 accidents in 1981, the formulas predicted 18.2 fatal accidents versus 13 actual and 31.3 injury accidents versus 42 actual. It should be noted that predicted severity values represent expected long-term annual rates and should be used with caution when estimating severity at individual crossings for a short time.

Results of the second evaluation of the severity formulas are based on the premise that, for accidents properly ranked by predicted severity, those at the top of the list (the most severe) should have a higher than average number of actual fatal and injury accidents. On the other hand, accidents at the top of a randomly ranked list should have only an average number of actual fatal and injury accidents. The ratio of actual accident severity for a set of accidents ranked by predicted severity to actual accident severity for the same size set of accidents ranked by random selection is a measure of the formula's ability to identify more severe accidents. This measure is referred to as the power factor for the prediction formula.

The power factors for the fatal and injury formulas for sets of accidents, ranked by predicted severity, are given in Table 5. The table indicates, for example, that for the top 100 ranked accidents the power factors for the fatal and injury formulas are 1.91 and 1.52, respectively. This means that the top 100 accidents ranked by the formulas have 1.91 and 1.52 times the number of fatal and injury accidents, respectively, as a randomly selected group of 100 accidents. Similar comparisons are made for the top 500 and 1,000 accidents. The results all show that the fatal and injury severity formulas are quite effective in predicting accident situations that tend to be more severe than the average.

TABLE 5 Ranking Performance of Severity Formulas

No. of Ranked Accidents	Fatal Severity Formula Power Factors[a]	Injury Severity Formula Power Factors[a]
100	1.91	1.52
500	2.24	1.24
1,000	2.13	1.26

[a] Actual severity for ranked group of accidents/actual severity for randomly selected group of accidents.

ACKNOWLEDGMENT

The original work on the resource allocation procedure and the present study were sponsored jointly by the FHWA Offices of Research, Development, and Technology and the FRA Office of Safety. The author expresses his appreciation to Janet Coleman, FHWA, and Bruce George, FRA, for their support and guidance and to Edwin Farr, Mary Cross, and Peter Mengert of TSC for their technical contributions to this study.

REFERENCES

1. E. Farr and J. Hitz. Accident Severity Prediction Formula for Rail-Highway Crossing. FHWA, U.S. Department of Transportation, July 1983. (Available from the Transportation Systems Center, Kendall Square, Cambridge, Mass. 02142.)
2. R. Coulombre, J. Poage, E. Farr, and J. Hitz. Summary of the Department of Transportation Rail-Highway Crossing Accident Prediction For-

mula and Resource Allocation Model. Report DOT-TSC-FRA-82-1. FRA, U.S. Department of Transportation, Sept. 1982.
3. J. Hitz and M. Cross. Rail-Highway Crossing Resource Allocation Procedure User's Guide. Report FHWA-IP-82-7. FRA, U.S. Department of Transportation, Dec. 1982.
4. J. Hitz, ed. Summary Statistics of the National Railroad-Highway Crossing Inventory for Public at Grade Crossings. Report FRA-RPD-78-20. FRA, U.S. Department of Transportation, Sept. 1978.
5. P. Mengert. Rail-Highway Crossing Hazard Prediction Research Results. Report FRA-RRS-80-02. U.S. Department of Transportation, March 1980.
6. E. Farr. Rail-Highway Crossing Resource Allocation Model. Report FRA-RRS-81-001. U.S. Department of Transportation, April 1981.
7. Rail-Highway Crossing Accident/Incident and Inventory Bulletin No. 4. FRA, U.S. Department of Transportation, 1981.
8. Risk Concepts in Dangerous Goods Transportation Regulations. National Transportation Safety Board, U.S. Department of Transportation, 1971.
9. D.R. Cox. Analysis of Binary Data. Halstead Press/John W. Stey, New York, 1970.
10. J. Coleman and G.R. Stewart. Investigation of Accident Data for Railroad-Highway Grade Crossing. In Transportation Research Record 611, TRB, National Research Council, Washington, D.C., 1976, pp. 60-67.

Publication of this paper sponsored by Committee on Railroad-Highway Grade Crossings.

Stability and Other Considerations in Simulation Analysis of Signal Control

FENG-BOR LIN

ABSTRACT

Lack of understanding of the nature of simulation and the characteristics of a system to be simulated can result in misuse of simulation models and simulation results. To promote better applications of simulation models to the evaluation of signal controls, three problems related to the generation and interpretation of simulation data are discussed in this paper. These problems include the stability of simulation results, the use of seed numbers for generating probabilistic events, and the aggregation of input data. Using simple examples of signal control, several fallacies in the application of signal simulation models are illustrated. Suggestions for avoiding these fallacious applications are presented.

Simulation models are increasingly used to aid in the design and evaluation of signal control systems. Some of these models, such as UTCS-1 (1) and NETSIM (2), are intended for general application in the evaluation of traffic control alternatives. These models require microscopic simulation of traffic flow characteristics to approximate the real world.

Experience with existing microscopic simulation models has produced a wealth of information on the potential and limitations of applying such models (3). Current concerns appear to focus on model enhancement, user needs and constraints, resource requirements for model application, and promotion and implementation of application by the traffic engineering community. The problem of experimental design for simulation analysis has also drawn some attention.

With increased use of simulation models for evaluation purposes, the risk of misuse and misinterpretation of simulation results can be expected to increase. A reason for this is that simulation models require substantial user interactions. An evaluation model is essentially a tool for data collection. Consequently, simulation results should be treated as a sample of observations. Estimates obtained from such a sample should be subjected to statistical tests for interpretation. It follows that experimental design should be an important part of simulation analysis. At issue is how, within the capability of a model, a user can apply the model efficiently to obtain statistically valid estimates.

The experimental design for simulation analysis is a profound subject. It requires a comprehensive understanding of the characteristics of a system to be simulated and the nature of simulation. At the present time, such an understanding is nonexistent. This is due in part to the large number of different systems a model has to accommodate. High costs and the reliability issue associated with the use of a model are also contributing factors. Nevertheless, there are several aspects of simulation application

that concern experimental design and can be readily discussed to benefit users of simulation models. These include the stability characteristics of simulation outputs, use of seed numbers for generating probabilistic events, and data aggregation. The purpose of this paper is to discuss the nature of these features.

STABILITY OF SIMULATION OUTPUTS

The operation of a signal control can be characterized by a set of measures of performance. These measures are in fact random variables because of the probabilistic nature of signal operation. A common purpose of simulation studies is to estimate the true value of a measure of performance. In this undertaking, users of a simulation model will have to deal with the stability of simulation outputs either directly or indirectly.

There are two primary features that determine the stability of simulation outputs. One is the dependence or independence of an estimate on the length of the simulation period. The other is the variation of estimates from true means. These features influence the length of a simulation run and the number of replicated runs needed to obtain reliable estimates. A discussion of these stability-related features based on vehicle delays follows.

Stable, Metastable, and Unstable States

Control performance can be classified in three states: stable state, metastable state, and unstable state. There are no clear-cut boundaries between stable state and metastable state nor between metastable state and unstable state. Nevertheless, these states have distinct characteristics that render them identifiable from simulation outputs.

A stable state usually exists when traffic volumes are light or moderate. Under this circumstance, the average delay of a traffic flow is governed primarily by the flow rate. For a given arrival pattern (e.g., random arrival), the sequence of the arriving headways or that of an event (e.g., gap acceptance) has little influence on the estimated mean value of a measure of performance. Using arriving headways as an example, this implies that rearranging the sequence of arriving headways will produce only slight changes in the performance of a control. In a simulation analysis, rearranging the sequence of arriving headways can be accomplished by using different seed numbers in replicated runs. It follows that estimated measures of performance from replicated runs will have only small variations. Another characteristic of a stable state is that estimated mean values of measures of performance can reach stabilized values as the simulation process is advanced. In other words, the estimates are time independent. An example of the stable state is the average delay of a flow of 400 vehicles per hour (vph) under a pretimed control shown in Figure 1.

In a metastable state, average delay depends not only on the flow rate but also on the sequence of arriving headways. A headway sequence of 3.2, 1.6, 5.3, ..., 10.2, 3.7, 2.1 sec, for example, can produce an estimate significantly different from that produced by a reversed sequence of 2.1, 3.7, 10.2, ..., 5.3, 1.6, 3.2 sec. Therefore, replicated simulation runs with different seed numbers can result in large variations in the estimated mean values of measures of performance. However, these estimates are still time independent (i.e., they can still reach stabilized values as the simulation process

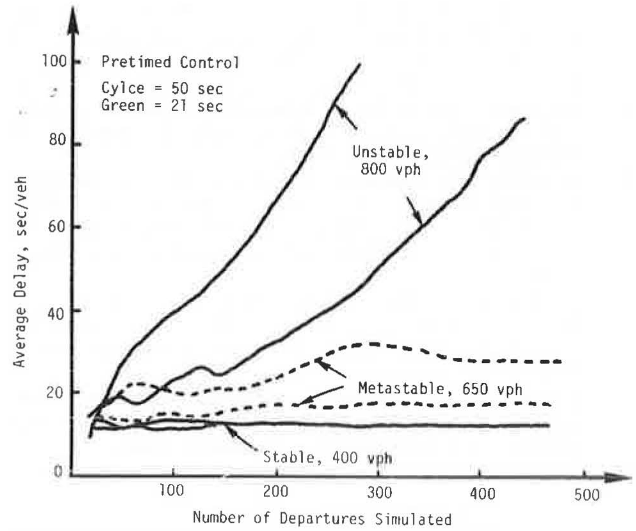

FIGURE 1 States of performance of signal controls.

continues). These characteristics are shown in Figure 1 by two simulation runs for a flow of 650 vph.

A metastable state may emerge when traffic flows are heavy enough to induce occasional carry-over of queuing vehicles from one cycle to the next. As the flows increase further, the queue length in a lane may grow with time if the same flows persist. Consequently, the performance of a control may become unstable. In an unstable state average delay depends not only on the flow rate and the sequence of arriving headways but also on the length of the simulation period. The longer the simulation period, the longer the average delay becomes. The estimated mean values of measures of performance are time dependent, and variations in the estimates from replicated runs can be expected to be large. The characteristics of the unstable state are also shown in Figure 1 by two replicated simulation runs.

The time-dependent features of a queuing system are rarely treated in the context of queuing theory because of mathematical complexities. In classic queuing theory, the operation of a system is usually assumed to be in a steady state. This implies that the average performance characteristics of a system do not change with time. This approach creates confusion when attempts are made to compare the output of a steady-state queuing model with simulation results. One example of such confusion involves Webster's delay formula (4). The approximate form of this formula is

$$D = 0.9\left\{[C(1-x)^2/2(1-xy)] + [y^2/2Q(1-y)]\right\} \quad (1)$$

where

D = average delay of vehicles in a traffic lane,
C = cycle length,
x = effective green-to-cycle length ratio,
y = saturation ratio, and
Q = flow rate.

This formula is a steady-state queuing model for a flow pattern with random arrivals and a uniform flow rate of Q. It assumes that the flow rate will persist indefinitely. As a result, when the saturation ratio y approaches 1.0 the estimated average delay approaches infinity. For most signal operations, this phenomenon cannot be observed in the

field because heavy flows do not last forever or even for a very long time. When heavy flows induce an unstable state, shorter simulation periods will result in shorter estimated average delays. This is shown in Figure 2. For this reason a comparison of simulation output with the output of a steady-state model becomes meaningless unless a very long simulation period is used. This also underscores the importance of determining whether simulation results are time dependent before they are used for comparative analyses.

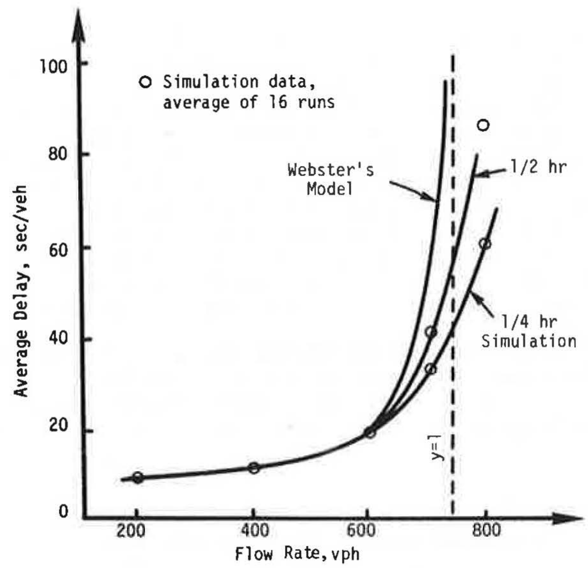

FIGURE 2 Comparison of simulated delays with Webster's estimates under a pretimed control.

To facilitate identification of this time-related stability feature, the simulation period beyond a nonrecording transient interval may be divided into time blocks of equal length. Each block should be at least equal to the expected maximum cycle length. Let x_{ij} represent the value of the jth observation (e.g., delay of a vehicle) in the ith time block. For each block the estimated mean $\bar{x}_i$ and standard deviation S_i can be determined from x_{ij} based on m_i observations. In addition, the grand mean $\bar{x}$ and the grand standard deviation S for all the time blocks simulated can be determined.

If variable flow rates are used as inputs to approximate an actual flow pattern, the simulation period should coincide with the actual duration of the flow pattern. In this case, simulation results may become time dependent for a short time and then stabilize. At the end of the simulation, the grand mean $\bar{x}$ should be printed out as a function of the number of time blocks simulated. This information can be used to determine whether and when the operation of a control is time dependent.

If uniform flow rates are used instead, users may be allowed to specify a minimum simulation period and a maximum simulation period beyond the transient interval. The minimum period is for the purpose of obtaining a sufficient number of observations before the stability check begins. Ten minutes of real time is probably a reasonable period to use. The maximum simulation period can be based on cost consideration. Between the minimum and the maximum, an algorithm incorporated into a simulation model may be used to determine whether simulation results have reached stabilized values or whether the operation of a control has become unstable.

A simple approach to detecting an unstable state is to examine the trend of $\bar{x}_i$ for several consecutive blocks (e.g., 4 or 5) from block a to block b. Block b represents the block that has just been simulated. If $\bar{x}_i$ increases monotonically from block a to block b, there is reason to suspect that an unstable state exists. In such a case, the estimate $\bar{x}_b$ of the current block may be compared with the estimate $\bar{x}_a$ several blocks earlier. The purpose is to determine whether $\bar{x}_b$ is significantly greater than $\bar{x}_a$ by a specified amount (e.g., 5 sec/vehicle). This can be carried out with a simple t-test.

In choosing a t-test, one should realize that the true standard deviations of the observations in two time blocks are unknown and are not necessarily the same. Therefore, a proper test is to determine the following statistics (5):

$$t = (\bar{x}_b - \bar{x}_a - \delta)/[(S_a^2/m_a) + (S_b^2/m_b)^{1/2}] \quad (2)$$

and

$$f \simeq [(S_a^2/m_a) + (S_b^2/m_b)]^2 / \{[(S_a^2/m_a)^2/(m_a - 1)] + [(S_b^2/m_b)^2/(m_b - 1)]\} \quad (3)$$

where

t = t-value of the difference $\bar{x}_b - \bar{x}_a$;
f = approximate degrees of freedom;
m_a, m_b = sample sizes of block a and block b, respectively;
S_a, S_b = sample standard deviations of block a and block b, respectively; and
δ = specified level of difference.

Given a level of significance α and degrees of freedom f, a critical t-value, denoted as t_c, can be estimated from a statistical table of t-distribution. If the t-value obtained from Equation 2 is greater than t_c, one can conclude that $\bar{x}_b$ differs significantly from $\bar{x}_a$ by the specified amount of δ. Consequently, one may assume that an unstable state exists.

This approach is not infallable. Figure 1 shows an example. One of the simulation runs for the metastable operation reveals that, between a total departure of about 140 to about 280 vph, the average delay increases monotonically. The equivalent time is about 12 min. The operation of the control in this period may be labeled unstable and the simulation process is subsequently terminated. If the process is allowed to continue, however, the estimate of the average delay stabilizes. Such an incorrect identification will likely occur only when the operation of a control falls in a gray area between a metastable state and an unstable state. Therefore, it will in fact force users to interpret the simulation results more cautiously.

Whether simulation results have reached stabilized values may be checked in a similar manner on the basis of a few decision rules. First, the block mean $\bar{x}_i$ should not display a monotonic increase or decrease from block a to block b. When this condition exists, the average value of the block means of block a through block b can be determined. Next, the proportion of observations having values greater than this average can be determined for each block. If the proportion is about 50 percent in each block

and it does not show a monotonic increase or decrease from block a to block b, the performance of the control may have stabilized. A more stringent test may then be performed to determine whether the proportion in block b is significantly different from that a block a. If the difference is not significant, the simulation results may be assumed to have reached stabilized values. The simulation process may be terminated if the estimated means of measures of performance have satisfied a desired level of accuracy.

If a model does not have an internal algorithm to check time-related stability, users will have to perform at least two replicated runs with different simulation periods. When an unstable state exists, the longer run will give a much higher estimate of average delay. In some cases, the number of runs may be reduced to one if the grand means at the end of each time block or the block means are printed out as a function of the number of blocks simulated. This information allows users to visually examine the performance of a control over time.

Within-Run and Between-Run Variations

In a simulation run, individual vehicle delays will vary. These within-run variations affect the accuracy of estimated average delays. For a given flow pattern, simulated average delays can also vary from one run to another. These between-run variations result from changes in operating conditions without a change in the average characteristics of a simulated flow pattern. Changes in the operating conditions can be induced by using different seed numbers in replicated runs for generating probabilistic events. For example, a seed number may lead to a headway sequence of 4.2, 1.5, ..., 10.2 sec. A new seed will certainly produce a different sequence.

In general, within-run variations can be expected to increase with flow rate. This is inherent to the nature of a queuing system. For example, assume that there is a single channel queuing system with random vehicle arrivals and random service time. Let λ represent the arrival rate and β the service rate. It can be shown (6) that, if λ and β remain unchanged, the cumulative probability density function for the waiting times is

$$f(w \leq t) = (\lambda/\beta) - [(\lambda/\beta)\, e^{-(\beta-\lambda)t}] \tag{4}$$

where $f(w \leq t)$ represents the probability of a waiting time w less than or equal to t.

Based on this function, the average waiting time becomes

$$\mu = \lambda/\beta(\beta-\lambda) \tag{5}$$

and the standard deviation of the waiting times is

$$\sigma = [2\lambda/\beta(\beta-\lambda)]^{1/2} \tag{6}$$

Equation 6 reveals that σ increases with the arrival rate λ. This characteristic is undesirable from the viewpoint of simulation. It means different signal control problems require different simulation periods in a single run to achieve a given level of accuracy in estimates. It also implies that heavier flows may need a much longer simulation period.

To determine whether simulation results have reached a specified level of accuracy, one has to have estimates of the true mean (e.g., μ) and the true standard deviation (e.g., σ) of a measure of performance. Let $\bar{x}$ and S be the estimates of mean and standard deviation, respectively. Also let n be the number of observations used in determining $\bar{x}$ and S. Regardless of the sample size n, the confidence interval of $\bar{x}$ can be determined as (5)

$$R = \bar{x} \pm t_c[S/(n-1)^{1/2}] \tag{7}$$

where

R = confidence interval of $\bar{x}$, and
t_c = critical t-value corresponding to n - 1 degrees of freedom and a level of significance α.

If the service rate β in Equation 6 is 400 vph, the value of σ given by the same equation equals 4.2 sec/vehicle for an arrival rate λ = 200 vph and 7.4 sec/vehicle for λ = 300 vph. At a level of significance α = 5 percent, the value of t_c approaches 1.96 for large n. Using the σ values for S in Equation 7, one can see that the sample size required to reduce the confidence interval to within 1 sec of the estimated mean is about n = 68 for λ = 200 vph and about n = 211 for λ = 300 vph. These sample sizes are equivalent to 0.4 hr of observations for λ = 200 vph and 0.7 hr for λ = 300 vph.

The between-run variations are smaller than the within-run variations because they involve the mean of each run. Nevertheless, such variations can be too large to ignore. Table 1 gives an example of the between-run variations in the average delays of vehicles. The vehicles are subjected to a two-phase fully actuated control that uses presence detectors. Each phase has two lanes. The flow rates are the same in all the lanes.

TABLE 1 Between-Run Variations in Average Delays Resulting from a Fully Actuated Control

Simulation Run No.	Average Delay (sec/veh) Lane Flow (vph) 200	400	600	700
1	4.0	9.6	20.1	49.7
2	5.1	10.1	22.3	59.6
3	4.8	9.7	23.1	55.6
4	4.3	11.3	22.9	52.3
5	5.2	9.8	28.3	64.5
6	4.6	8.8	22.4	51.3
7	5.2	9.1	22.4	69.9
8	4.7	9.0	22.3	53.8
9	4.2	10.1	23.1	63.5
10	4.9	9.9	22.6	50.2
$\bar{x}$	4.7	9.7	23.0	57.0

Note: Simulation period = 0.5 hr; extension interval = 0 sec.

When the major phase flows are under 400 vph/lane, the table shows that any of the 10 runs results in an estimated mean within about 1 sec/vehicle of the overall mean $\bar{x}$. At higher flow rates, estimated means of individual runs begin to differ more and more from the overall mean. This example reveals the need to perform replicated runs and the potential risk of using the results of a single run for comparative analysis.

When replicated runs are performed, the estimation of the mean and standard deviation of a measure of performance becomes more complicated. The analysis of variance techniques described by Brownlee (5) may be used under the circumstance. These techniques require that the individual observations (e.g.,

vehicle delays) obtained in each run be stored in one way or another as a data file. The results of several runs would form a matrix of individual vehicle delays for analysis.

Obviously, the need to treat and analyze such a matrix after each additional run would be a heavy burden on most users. A more practical approach is to obtain the same (or approximately the same) number of observations in each run. As an approximation, this can be accomplished by using equal simulation periods for all the replicated runs. In this case, users have to deal only with the estimated mean of each run. Let $\bar{x}_r$ represent the mean estimated from the rth replicated run and k the number of runs. Then, the true mean can be estimated as

$$\bar{x} = \sum_{r=1}^{k} \bar{x}_r/k \tag{8}$$

and the standard deviation of the estimated mean, denoted as $S_{\bar{x}}$, becomes

$$S_{\bar{x}} = \left\{ \left[\sum_{r=1}^{k} (\bar{x}_r - \bar{x})^2 \right] /k(k-1) \right\}^{1/2} \tag{9}$$

For runs with equal numbers of observations, it can be shown analytically that $\bar{x}$ and $S_{\bar{x}}$ are the same as those obtainable from a more complicated analysis of variance technique. These estimates can be used in several ways. For example, they allow users to determine the confidence interval of $\bar{x}$. In this application the term $S/(n - 1)^{1/2}$ in Equation 7 can be replaced by $S_{\bar{x}}$, and the critical t-value (i.e., t_c) has k - 1 degrees of freedom. For comparative analysis of estimated $\bar{x}$, it is advisable that the confidence intervals of these estimates be approximately the same. Otherwise one would be comparing estimates of different levels of accuracy. Given this understanding, Equations 2 and 3 can be modified to test the significance of the difference between the operations of alternative controls as represented by $\bar{x}$. For this purpose, δ of Equation 2 can be set to zero; $S_a{}^2/m_a$ and $S_b{}^2/m_b$ are equivalent to the values of $S_{\bar{x}}$ of two alternative controls; and m_a and m_b represent, respectively, the numbers of replicated runs performed for the alternatives.

So far no reliable methods are available to introduce between-run variations into a single run and thus eliminate the need to perform replicated runs. To reduce the number of replicated runs, one will have to find a way to rapidly obtain narrow confidence intervals of $\bar{x}$. This can be achieved only if one can produce a large reduction in $S_{\bar{x}}$ from each additional run. One approach to this problem is to use negatively correlated random numbers in replicated runs (7). This approach can be easily implemented by using a random number R ($0 \leq R \leq 1.0$) in one run and 1^{-R} in another.

USE OF SEED NUMBERS

Ideally, separate events (e.g., headways in a lane or lane changes) in a simulation run should be generated from different seed numbers. This is not difficult to do because a single seed number can be used to generate randomly a string of seed numbers. These seed numbers can then be used to generate different events. In the simulation of a large system, however, numerous events are encountered. If an event is assigned a separate seed number, memory storage requirements may be increased substantially. Central processing unit time will also increase. To simplify the task, a simulation model may rely on a single string of random numbers generated from a seed number to simulate various events. Questions have been raised about the desirability of such a practice (3).

The use of a single string of random numbers for different events will certainly destroy the ability of a simulation model to perform controlled analyses. This is indeed undesirable if one is to compare alternative controls based on a single run for each alternative. However, one should also realize that in such comparative analyses the only event that a simulation model can control precisely is input flow patterns in the form of arriving headways. When vehicles are processed downstream and subjected to influences by control strategies and by conflicts between traffic flows, not many events can remain the same under alternative controls. Therefore, the real issue is whether a single string of random numbers can produce distributions of various events that conform to predetermined distributions.

To provide an insight to this problem, consider the generation of headways for vehicles in a traffic lane. With a random arrival pattern, the headways follow a shifted negative exponential function:

$$f(h \geq t) = e^{[(t-\tau)/(T-\tau)]} \tag{10}$$

where

$f(h \geq t)$ = probability that a headway h is greater than or equal to t;
T = average headway of vehicles; and
τ = minimum headway, taken to be 1 sec.

Assume that the arrival of vehicles in that lane is only one of 15 events to be generated with a single string of random numbers. An event may not have any sampling unit for observation during a particular period of time in the simulation process. Therefore, in one case two or more successive random numbers may be used consecutively for generating headways. In another, 14 successive random numbers may be used to generate other events before the next one is used again for the headways. Of the 80 distributions of headways randomly generated in this manner for a flow rate of 600 vph, 3 do not conform to the shifted negative exponential function at a level of significance of 5 percent (chi-square test). Another 4 have mean headways significantly different from the intended mean of 6 sec. Therefore, there is approximately a 9 percent chance that the use of single-string random numbers will fail to produce the desired arrival pattern for a flow of 600 vph. A similar experiment for a flow of 200 vph also results in a failure rate of 9 percent. In contrast, when headways are determined sequentially from every random number generated, the failure rate is 6 percent for a flow of 600 vph and 9 percent for a flow of 200 vph.

These brief analyses show that, from the viewpoint of generating predetermined distributions of events, the use of a single string of random numbers is as good as the use of multiple strings. The analyses also reveal that a generated distribution may be quite different from what is desired. The results of a simulation analysis may be made more reliable if seed numbers are chosen carefully to ensure that generated distributions conform to intended distributions.

AGGREGATION OF INPUT DATA

The use of aggregated data results in a loss of useful information. For simulation analysis of signal controls, it may prompt a model to produce estimates quite different from those obtained from the use of disaggregated data.

One type of data aggregation commonly encountered in the use of a signal simulation model is average flow rates. The impact of such practice on estimated average vehicle speed has been examined in two previous studies based on UTCS-1 and TRANSYT (8-10). The time aggregation of flow rates was found to have insignificant impact on estimated average vehicle speeds.

It should be cautioned, however, that different measures of performance may have different sensitivities to the aggregation of input data. Table 2 gives an example of this possibility in terms of average delays resulting from a two-phase, fully actuated control. This control employs 50-ft-long presence detectors. The extension intervals are set at 1.5 sec and the maximum greens at 50 sec. Each phase has two lanes with equal traffic volumes. The actual flow rates per 5-min interval over a 1-hr period are assumed to have the following relative values: 0.85, 0.90, 0.95, 1.0, 1.1, 1.2, 1.2, 1.1, 1.0, 0.95, 0.90, and 0.85. The average flow rate for the entire hour has a relative value of 1.0.

TABLE 2 Effects of Time Aggregation of Flow Rates on Estimated Average Delays

	Average Delay (sec/veh)			
	Uniform Rate Simulation		Variable Rate Simulation	
	vph/Lane		vph/Lane	
Run No.	400	600	400	600
1	10.7	20.6	11.3	20.4
2	11.0	22.1	11.6	27.6
3	9.5	21.7	10.1	24.3
4	10.6	20.0	10.9	21.2
5	10.9	20.2	11.5	22.5
6	9.6	21.0	9.9	26.7
7	10.1	19.3	10.5	20.0
8	9.6	21.3	10.0	23.5
9	9.8	21.0	10.5	21.7
10	10.5	22.2	10.9	28.9
$\bar{x}$	10.2	20.9	10.7	23.7
S	0.6	0.9	0.6	2.8

Note: Simulation period = 1 hr; extension interval = 1.5 sec.

The performance of the control is simulated in two ways for average flow rates of 400 and 600 vph. One is based on the variable flow rates and the other uses the average flow rates as input. In both cases, arriving headways are assumed to distribute according to the shifted negative exponential function of Equation 10.

The results given in this table reveal that the use of average flow rate leads to smaller between-run standard deviations. This is a desirable feature if the estimated average delays are insensitive to the data aggregation. Unfortunately, statistical tests based on Equations 2 and 3 show that data aggregation tends to produce lower estimated averages. The differences are significant at a level of significance of 5 percent for flow rates of 600 vph. The difference is insignificant for a flow of 400 vph. Thus, data aggregation may significantly bias estimates under heavier flow conditions.

This simple example shows the potential risks of using average flow rates. However, one should not conclude that average flow rates should not be used for simulation. What is important is that the use of simulation results should be consistent with the simulation conditions. The purpose of a simulation analysis should dictate the choice of the input data. For simulating the actual operation of a signal control for the purpose of evaluation, it is advisable not to use average flow rates. For comparative analyses of alternative controls, uniform as well as variable flow rates may be used if it is known that the data aggregation will not alter the estimated relative merits of the controls.

Another common practice of data aggregation is the use of average headways. For example, the average discharge headway of queuing vehicles in a given queuing position may be used in place of a probability distribution of individual headways. This type of data aggregation can provide adequate estimates under a variety of circumstances, but there is always a possibility that it may produce a distorted picture of signal operation.

Consider the operation of a fully actuated control based on presence detectors. Under this mode of control a vehicle can extend the green if it actuates a detector before a vehicle ahead departs from the detection area and the extension interval expires. The arrival time of a vehicle at the upstream edge of a detector and the departure time from the downstream edge of the same detector can be measured in the field. Table 3 gives the average values of these flow characteristics in relation to a 50-ft detector as observed on Almond Street, Syracuse, New York. These averages indicate that a vehicle in the third queuing position waiting upstream of the detector at the onset of the green needs an average of 5.2 sec to reach the detector. The vehicle ahead in the second queuing position requires 5.6 sec to depart from the detection area. Vehicles in the queuing positions farther upstream have average arrival times longer than the average departure times of the vehicles immediately ahead.

TABLE 3 Average Arrival and Departure Times of Queuing Vehicles in Relation to a 50-ft Detector

Queue Position	Departure Time (sec)	Arrival Time (sec)
1	3.1	-
2	5.6	-
3	7.9	5.2
4	10.1	8.1
5	12.2	10.6
6	14.4	12.9
7	16.4	15.1
8	18.5	17.2

Note: Data gathered for Almond Street, Syracuse, New York.

Based on the average arrival times and departure times, one can conclude that, if no extension is given to the green after each departure and only one lane flow is associated with a signal phase, vehicles in the back of a long queue will face a certain premature termination of the green. In reality, arrival times and departure times are probabilistic. For example, the observed departure times of vehicles in the first queuing position range from 1.2 to 5.8 sec. The field data show that under the same conditions vehicles in the fourth queuing position have a 10 percent chance of facing a premature termination of the green. For those in the seventh queuing position the chance is 40 percent.

The effect of using the averages in estimating average delays is shown in Figure 3. This figure is based on a two-phase control with two lanes in each phase. The lane flow Q in the major phase is twice as heavy as that in the minor phase. The extension

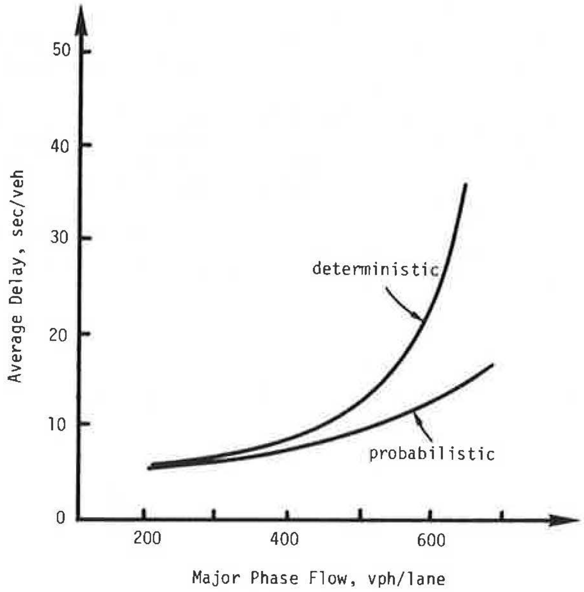

FIGURE 3 Effects of using average headways (simulation period = 1 hr).

interval is 0 sec and each phase has a clearance interval of 4 sec. The signal is allowed to rest in red if no actuations of the detectors take place. To extend the green continuously under these conditions, at least one vehicle should always be in one of the detection areas associated with a given signal phase. And, when the traffic flows are light, the delay of a vehicle is primarily attributable to the need to decelerate. The curve labeled probabilistic is obtained from the actual probabilistic distributions of the arrival and the departure times. The other curve labeled deterministic is based on the averages given in Table 3.

When the flows are light ($Q \leq 400$ vph), queue lengths are short and the opportunity for premature termination of the green rarely exists. As a result, the use of averages brings about only slightly higher estimates (up to 1.5 sec/vehicle). When the flows exceed 500 vph, the averages given in Table 3 overestimate the probability of premature termination of the green. The resulting estimates of average delays deviate significantly from the probabilistic estimates.

CONCLUSIONS

The stability characteristics of simulation results can be assessed in terms of (a) their dependence on or independence of the length of simulation period, (b) within-run variation, and (c) between-run variation. These characteristics are inherent in the operation of a signal control, but they complicate the application of simulation models.

It is convenient to infer from simulation results that one alternative improves the efficiency of a control by a certain percentage. But it is more difficult to determine whether the alleged improvement is real. For proper interpretation of simulation results, a user first has to know whether the results are time dependent. Furthermore, when uniform flow rates are used as inputs and the operation of a control is not in an unstable state, a user has to ensure that the results from a simulation run represent stabilized estimates of the performance of a control.

A simple algorithm may be incorporated in a model to detect the existence of an unstable-state operation. The algorithm should also be capable of determining whether estimates of measures of performance have reached stabilized values. Such an algorithm should at least be implemented for system-wide estimates and, at the option of a user, for estimates related to individual components of a system. To assist users in the interpretation and comparison of simulation results, a model should provide the following outputs for key measures of performance: estimated mean, standard deviation of the mean, and number of observations. Estimated means recorded as a function of elapsed real time are also useful.

Between-run variations always exist in microscopic simulation of signal control. Because of this, replicated runs are necessary unless the accuracy of simulation results is not a concern. Introducing between-run variations into a single run can eliminate the need for replicated runs. But no established mechanisms are available for such an application.

When replicated runs are performed, it is advisable to obtain an equal number of observations for each run. This will simplify the estimation of the mean of a measure of performance and its standard deviation. Negatively correlated simulation runs are preferred to independent runs. They can reduce the number of runs needed to achieve a specific level of estimation accuracy.

The use of a single string of random numbers would make comparative analysis based on identical sequences of events impossible. One should realize, however, that in a simulation analysis only arriving headways can really be rigidly controlled for comparisons of alternatives. Other events resulting from the interaction of flows and signals will vary from one alternative to another. Only the probability distributions of such events can be controlled through the use of multiple strings of random numbers. In generating the distribution of an event, the use of multiple strings does not appear to have a real advantage over the use of a single string. In fact, both approaches may fail to generate a desired distribution. Therefore, it may be more important to ensure that the use of a seed number will generate a distribution correctly.

Aggregation of data to provide inputs runs the risk of incurring biased estimates. The bias may result from reduced variations in representing the real world. It may also result from incorrect representation of the operation of a control. Whenever possible, it is wise to avoid the use of aggregated data, particularly when heavy flows are involved.

The discussion presented herein perhaps overemphasizes the need to determine the level of accuracy of estimated measures of performance. The profound problems associated with the application of simulation models may be understated. The accuracy check of simulation results is meaningless unless a model is capable of generating reliable information. Furthermore, the level of accuracy of estimates has only to be compatible with the purpose of a simulation study. In any case, interpretation of simulation results should be done very cautiously if their accuracy is uncertain.

REFERENCES

1. Peat, Marwick, Mitchell, and Co. Network Flow Simulation for Urban Traffic Control System--Phase II: Vol. 4, User's Manual for UTCS-1 Network Simulation Model. FHWA-RD-73-86. FHWA, U.S. Department of Transportation, May 1973.

2. Traffic Network Analysis with NETSIM, A User Guide. FHWA-IP-80-3. FHWA, U.S. Department of Transportation, Jan. 1980.
3. Application of Traffic Simulation Models. TRB Special Report 194. TRB, National Research Council, Washington, D.C., 1981, 114 pp.
4. F.V. Webster. Traffic Signal Settings. Road Research Technical Paper No. 39. Great Britain Road Research Laboratory, London, England, 1958.
5. K.A. Brownlee. Statistical Theory and Methodology in Science and Engineering. Wiley, New York, 1965.
6. D.R. Drew. Traffic Flow Theory and Control. McGraw-Hill, New York, 1968.
7. J.R. Emshoff and R.L. Sisson. Design and Use of Computer Simulation Models. MacMillan, New York, 1970.
8. S. Yager and E.R. Case. Summary Evaluation of UTCS-1/NETSIM in Toronto. In TRB Special Report 194: Application of Traffic Simulation Models, TRB, National Research Council, Washington D.C., 1981, pp. 95-99.
9. S. Yager and E.R. Case. Using TRANSYT for Evaluation. In TRB Special Report 194: Application of Traffic Simulation Models, TRB, National Research Council, Washington, D.C., 1981, pp. 100-103.
10. D.I. Robertson. TRANSYT: A Traffic Network Study Tool. RRL Report LR253. Road Research Laboratory, London, England, 1969.

Publication of this paper sponsored by Committee on Traffic Control Devices.

Evaluation of Engineering Factors Affecting Traffic Signal Change Interval

MYUNG-SOON CHANG, CARROLL J. MESSER, and ALBERTO SANTIAGO

ABSTRACT

Engineering factors affecting traffic signal change interval (yellow and all-red) are reviewed, particularly in terms of drivers' perception and brake reaction time (t), and deceleration rate (d). Using driver behavior data collected from time-lapse cameras, two hypotheses were tested: (a) t and d are dependent on speed and (b) there is an interactive effect of t (prebraking) and d (postbraking) on drivers' braking distance. All hypotheses from statistical analysis results were accepted. It is concluded that joint consideration instead of independent consideration should be given to t and d when selecting their values. Furthermore, it is recommended that different t's and d's for different approach speeds should be used rather than a single value (as used in current practice) for all approach speeds in signal change interval design.

To stop or not to stop? That is the question asked when drivers see a green light ahead change to yellow and then to red. If the driver fails to respond safely, a major right-angle collision at the intersection is possible. On the other hand, if the driver overreacts, a hazardous rear-end collision is likely. Because of the complexity of the driver-vehicle-traffic control system involved and the potential severe consequences of system failure (a fatal accident), the design of the traffic signal yellow time and any following all-red interval should be optimized based on the best understanding of the engineering factors involved. The magnitude of the problem requires that traffic engineers do no less.

PERCEPTION AND BRAKE REACTION TIME (t)

AASHO (1) recommends a total of 2.5 sec for perception and brake reaction time. The Institute of Traffic Engineers (ITE) Handbook (2,3) assumed a perception-brake reaction time for a yellow signal of 1 sec. Actual drivers' stopping distance data reported by Williams (4) and Sheffi (5) are analyzed using different deceleration rates from 8 to 15 feet per second per second (fps²). The results, which indicate three categories of t under normal driving behavior, are as follows:

1. Forced stopping: When more than 85 percent of the drivers go through the intersection, those 15 percent or less of the drivers who decide to stop take less than 1 sec of perception and brake reaction time.
2. Indecision stopping: When half of the drivers decide to stop, the perception and brake reaction time is 1 to 1.5 sec.
3. Comfortable stopping: When the majority of drivers decides to stop, their perception and brake reaction time is 1.5 to 3.0 sec.

An analysis of Williams' (4) and Sheffi's (5) results also indicates that perception and brake

reaction time at high speeds may be less than that at low speeds. It also appears that, when a driver's position is farther away from the intersection, the perception and brake reaction time to the yellow onset may be longer than for drivers closer to the intersection. In addition, at a given speed, less variation in time t is presumed to exist for drivers closer to the intersection than for those farther away. These developments are based on drivers' physical limitations and lag time, which appears to fall between perception and brake reaction time.

DECELERATION RATE (d)

The ITE Handbook in 1976 (2) used 15 fps² as the deceleration rate. Bissell and Warren (6) and Parsonson and Santiago (7) suggested 10 fps² for the deceleration rate. AASHO (1) recommended 18.0 fps² at design speed of 60 mph and 20.0 fps² at design speed of 30 mph for dry pavements. For wet pavements, it recommended 9.7 fps² at 60 mph and 11.6 fps² at 30 mph, respectively.

It appears that the deceleration rates used by the ITE Handbook in 1976 and by AASHO for dry pavements are unrealistically high. Using data from Williams (4) and Sheffi (5), the required deceleration rates were analyzed using a perception-brake reaction time of 1 sec. The analysis indicates that the majority of drivers who stopped experienced deceleration rates of 5 to 11 fps² for approach speeds of from 25 to 55 mph. Only in the case of a few forced stops was a deceleration rate of 15 fps² observed. Furthermore, the results indicate the presence of different deceleration rates for different approach speeds. It is noted that the ITE Handbook in 1982 (3) suggested, without acknowledgment, the use of 10 fps² as the deceleration rate.

STUDY OBJECTIVE

It is noted that no previous study hypothesized the possible pair of t and d for different approach speeds. One value of t and d is used for all approach speeds. Further, no study examined the potential interaction of t and d. In other words, it can be hypothesized that braking distance and subsequent d are affected by t.

In this paper these problems are addressed using data collected at one intersection in College Station, Texas. Specifically, the objectives of this paper are to examine whether t and d are dependent on speed, and if there is an interactive effect of t and d on drivers' selected braking distance.

DATA COLLECTION

Two time-lapse cameras were used to collect data. Each camera was mounted, using a ladder up a utility pole, on one approach of each street. Both time-lapse cameras were connected to remote control units that could be operated from inconspicuous locations. Both remote cameras were operated approximately 3 sec before yellow and continued operating until the first stopping car in each lane had made a complete stop. The film was shot at nine frames per second. Data collection activity specifically covered peak, off-peak, nighttime, and wet weather conditions. The study collected operational and physical data necessary to evaluate the engineering factors and driver responses affecting change interval design. The perception and brake reaction time is measured as the time elapsed from the onset of yellow until the brake light is observed to come on. Thus t will also include response lag time, which falls between perception and brake reaction.

BOUNDARY AND POPULATION VEHICLES

The vehicles within the observation distance from the stop line to the upstream outer boundary at the onset of yellow are the "boundary vehicles." All vehicles beyond the stop line and preceeding the upstream outer boundary at the onset of yellow are ignored. Further, vehicles turning right are excluded because their response is anticipatory and not a random response.

Among boundary vehicles, those vehicles that go through the intersection and those first vehicles stopped in each lane are the "population vehicles." Vehicles stopping second in each lane are excluded because their behavior is constrained by the first vehicle stopped in each lane (i.e., their probability of stopping = 1). Further, within the sample vehicles stopped in each lane, motorists braking during green through yellow, or braking coincident with the yellow onset, should also be excluded because their behavior is not a random response to the yellow signal but anticipatory or coincidental.

TEST OF HYPOTHESES

In this presentation, the coefficients of variables found from models are intentionally omitted and only graphic results are presented. Because they represent only effects based on mean values and their misuse may cause significant consequences, the coefficients will be presented in a subsequent paper after more data have been collected. It is noted that no high correlation was found between independent variables throughout the models presented.

Hypothesis 1

Drivers' perception and brake reaction time is affected by approach speed (V), distance at yellow onset (D), and the interaction of these two.

The best model obtained from stepwise regression procedure revealed that a driver's perception and brake reaction time is affected by approach speed, distance from the intersection at yellow onset, and the interaction of approach speed and distance. Specifically, it was found that

$$t = f(-v, D, -V \times d, V^2, D^2) \qquad (1)$$

The model and all variables were significant at $\alpha = 0.01$ and the R^2 was 0.47. Although $R^2 = 0.47$ is not overwhelmingly high, the model confirmed the fundamental factors affecting drivers' perception and brake reaction time.

Graphic presentation of the model (Figure 1) shows that a driver's perception and brake reaction time decreases as approach speed increases. Note that the approach speed is derived during green just before yellow onset. For the cases of high t the speed between yellow onset and brake actuation should be discounted because of coasting.

Hypothesis 2

Drivers' deceleration rates are affected by approach speed, distance from intersection at yellow onset, and perception and brake reaction time.

Similarly, stepwise regression was used and the

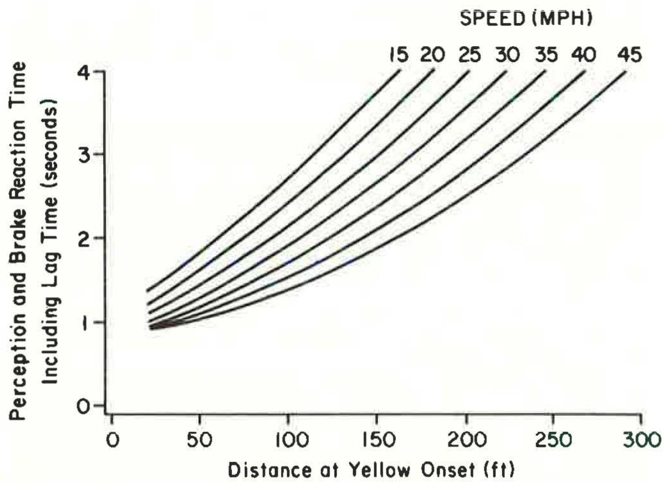

FIGURE 1 Perception and brake reaction time as a function of approach speed and distance at yellow onset.

model revealed that deceleration rate is a function of approach speed, distance at yellow onset, and interaction of distance at yellow onset and t. Specifically, it was found that

$$d = f(V, D \times t, -D^2) \quad (2)$$

The R^2 was 0.08 and the model was significant at α = 0.01. The interaction of D and t was significant at α = 0.05, and other variables were significant at α = 0.01. Even if the model is significant, no definite conclusion may be drawn from the data because of a weak relationship. However, the model is noteworthy and will be retested as more data become available.

Hypothesis 3

Braking distance is affected by approach speed and perception and brake reaction time. The interaction of prebraking (t) and postbraking (d) affects braking distance.

The best model found from stepwise regression procedure revealed that a driver's braking distance (B) is a function of the following variables:

$$B = f(V, t \times d, -V^2, -t^2, - d^2) \quad (3)$$

The R^2 was 0.51 and the model was significant at α = 0.01. The interaction of t and d was significant at α = 0.10 and the square of t was significant at α = 0.05. Other variables were significant at α = 0 01.

Graphic presentation of the model (Figure 2) indicates that

1. Drivers' braking distance increases as approach speed increases.
2. Drivers' braking distance decreases quadratically as deceleration rate increases. It is noted that the law of motion stating that braking distance is inversly proportional to deceleration rates is not applicable here because of the influence of t.
3. At low deceleration rates and at relatively farther distances from the intersection, slow-reacting drivers experienced the same comfortable deceleration rate as quick-reacting drivers even if their braking distances were shorter than those of the quick-reacting drivers. This is probably due to the speed reduction from coasting during the longer t periods of slow-reacting drivers.

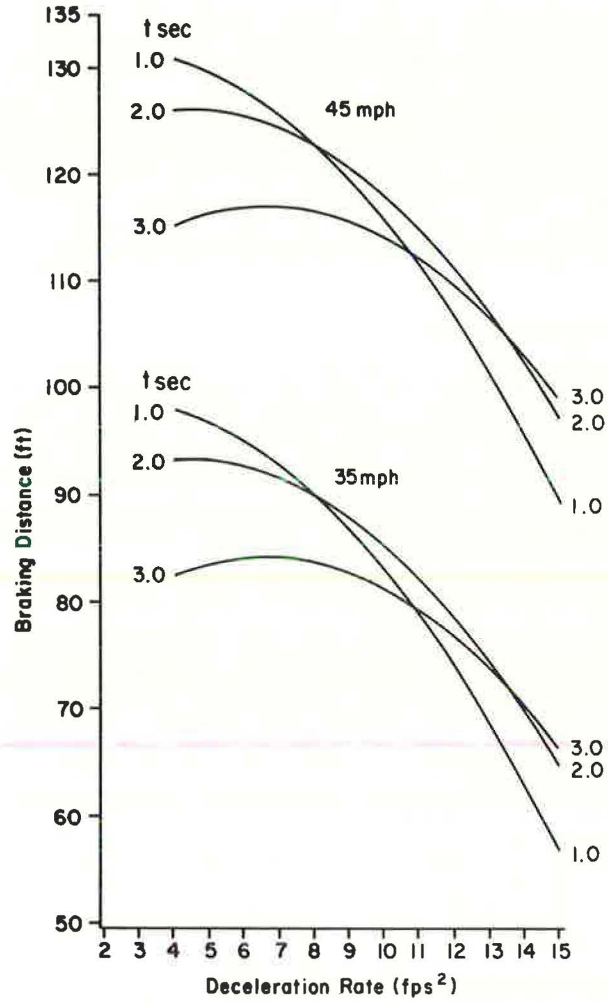

FIGURE 2 Breaking distance as a function of deceleration rate, approach speed, and perception and brake reaction time.

4. At high deceleration rates and at relatively closer distances from the intersection, the slow-reacting drivers experienced the same high deceleration rates as the quick-reacting drivers even though their braking distances were longer than those of the quick-reacting drivers. This is probably due to the human psychology of reducing accident risk by applying high deceleration rates very close to the intersection to compensate for slow reactions.

FURTHER STUDY

The Texas Transportation Institute (TTI), Texas A&M University, is analyzing more data obtained from six other sites. The study will address questions, including the probability of stopping or entering as a function of factors, involving signal change interval design.

CONCLUSIONS

The following conclusions were drawn from the data collected and apply to the experience with local intersections.

1. There was high variability in drivers' perception and brake reaction time (t). Higher t was observed not only for vehicles farther from the

intersection but also for vehicles closer to the intersection. Seventy-five percent of the drivers experienced 2 sec or less of t.

2. Median deceleration rates were found to be 12.5 and 10.5 fps^2 for each approach. The 85 percentile deceleration rates in descending order were 10.0 and 8.0 fps^2 for each approach.

3. Eighty-five percent of the drivers at the site applied the brakes within 240 ft of the intersection stop line.

4. Perception and brake reaction time is affected by approach speed, distance from the intersection at the onset of yellow, and the interaction of these two. Specifically, t decreases as approach speed increases at a given distance at yellow onset.

5. The deceleration rate appears to be affected by approach speed, distance at yellow onset, and perception and brake reaction time.

6. Braking distance is affected by approach speed, perception and brake reaction time, deceleration rate, and the interaction of t and d. Because braking distance is affected by the interaction of prebraking (t factor) and postbraking (d factor), joint consideration should be given the two in selecting the pair of t and d.

7. Because of conclusions 4-6, a different pair of t and d values for different approach speeds should be considered in designing change interval duration.

ACKNOWLEDGMENT

This paper is based on a pilot study conducted for the FHWA, U.S. Department of Transportation. The contents are the authors' and are not necessarily those of the FHWA.

REFERENCES

1. A Policy on Geometric Design of Rural Highways. AASHO, Washington, D.C., 1965.
2. Transportation and Traffic Engineering Handbook. Institute of Traffic and Transportation Engineers, Arlington, Va., 1976.
3. Traffic and Transportation Engineering Handbook. Institute of Traffic and Transportation Engineers, Washington, D.C., 1982.
4. W.L. Williams. Driver Behavior During the Yellow Interval. In Transportation Research Record 644, TRB, National Research Council, Washington, D.C., 1977, pp. 75-78.
5. Y. Sheffi and H. Mahamssani. A Model of Driver Behavior at High Speed Signalized Intersections. Transportation Science, Feb. 1981.
6. H.H. Bissell and D.L. Warren. The Yellow Signal Is Not a Clearance Signal. Institute of Transportation Engineers Journal, Feb. 1981, pp. 14-17.
7. P.S. Parsonson and A. Santiago. Design Standards for Timing the Traffic-Signal Clearance Period Must be Improved to Avoid Liability. In Compendium of Technical Papers, Institute of Traffic and Transportation Engineers, 1980.

Publication of this paper sponsored by Committee on Traffic Control Devices.

Abridgment

Strobe-Supplemented Red Signal Indications

TIMOTHY A. RYAN

ABSTRACT

A strobe-supplemented red indication (SSRI) is a red indication of a vehicle traffic control signal augmented by a flashing white light. The purpose of this project was to investigate the use of SSRIs. It was thought that SSRIs might be useful in reducing accident frequency at given locations. It was also through that those jurisdictions that have used SSRIs might have gained some valuable insights from their experiences with SSRIs. Thus, an investigation of these points was undertaken. The results of the project may be summarized as follows. (a) SSRIs have been used in a number of jurisdictions. The confirmed users are the city of Tampa, Florida; Metropolitan Dade County, Florida; the city of Portland, Oregon; Montgomery Country, Maryland; the city of Charleston, South Carolina; the state of North Carolina; the state of Maryland; the state of Kansas; the state of New York; and the District of Columbia. There may well be other users that were not identified in this project. (b) SSRIs have been used principally as an attempt to reduce excessive accident rates and to draw driver attention to an unexpected signal. (c) In most cases, the number of accidents and accident rates decreased following SSRI installation. Little statistical significance can be attached to these decreases. However, the lack of statistical significance may be due, at least in part, to the limited amount of data available. (d) In some jurisdictions it is believed that SSRIs have been effective, but officials in others are uncertain. There is no clear consensus on this question.

A strobe-supplemented red indication (SSRI) is a red indication of a vehicle traffic control signal augmented by a flashing white light. The purpose of this project was to investigate the use of SSRIs.

It was thought that SSRIs might be useful in reducing accident frequency at given locations. It was also thought that those jurisdictions that have used SSRIs might have gained some valuable insights from their experiences with SSRIs. Thus, an investigation of these points was undertaken.

Specifically, this effort was directed toward answering the following questions.

1. In which jurisdictions have SSRIs been used?
2. Why have SSRIs been used in those jurisdictions?
3. How have accident rates changed since SSRIs were installed?
4. Do officials in the jurisdictions in question believe that SSRIs have been effective?

LITERATURE SURVEY

SSRIs have been in use for at least 10 years and quite possibly longer. During that time a number of jurisdictions have experimented with SSRIs, and some continue to use SSRIs. However, little has been written about SSRIs. The information that is available on the topic comes primarily from unpublished reports or memoranda produced by jurisdictions that have used SSRIs [R.G. Edwards, Before and After Study, Strobe Lights, Route 17--Big Flats (Memorandum), New York State Department of Transportation, July 1980; T.G. Swenson, Final Report on Sg-23t-Use of Strobe Light in Traffic Signals, Kansas Department of Transportation; and L.R. Weatherby, The Effect of Raeco's Halo Light on Right-Angle Accidents at Two Intersections in Tampa, Florida, city of Tampa, Florida, June 1975]. Because these reports and memoranda were used as part of the data base for this project, the information in them is not discussed here.

SSRI USE

Hardware

One type of strobe unit is called a Halo by its manufacturer. The Halo is a circular ring that surrounds a normal 8-in. red traffic signal lens. A second type of strobe is called a Barlo by the manufacturer; the Barlo consists of a bar that is positioned in front of a red lens. A third type of strobe was identified by one of the responding jurisdictions. This type was a "vertical slit in center of lens." No further information about this type of unit was found.

Official Sanctioning of SSRI Experimentation

The Manual on Uniform Traffic Control Devices (MUTCD) (1) does not mention SSRIs. Thus, there is no official sanctioning of general use of SSRIs. However, a number of jurisdictions have received permission to experiment with SSRIs from the National Committee on Uniform Traffic Control Devices (NCUTCD) or the FHWA. Among the jurisdictions receiving such permission are the District of Columbia; Montgomery County, Maryland; the city of Portland, Oregon; the city of Charleston, South Carolina; the state of Kansas; and the state of Indiana. Of course, it is also quite possible that some jurisdictions may have installed SSRIs on their own initiative without seeking official permission.

Current Status of SSRIs

The most recent official statements regarding SSRIs appeared in the Federal Register on February 4, 1982, and January 10, 1983. The February 4, 1982, entry stated that a request to change the MUTCD had been made (Request IV-15), but that no change was being proposed. The following reason was given.

> Although results of the experimental projects have shown a significant reduction in accidents, the projects have been limited to single intersections. There is a need for further evaluation of the strobe device and the development of specifications for its design before its standard use might be proposed (p. 5254).

The January 10, 1983, entry simply stated that the "no change" response to Request IV-15 had been adopted as a final rule, "... pending availability of additional research or study data" (p. 1051).

PROJECT METHODOLOGY

Identification of Jurisdictions that Have Used SSRIs

The first list of jurisdictions that have used SSRIs was compiled from the author's own experience, a conversation with the manufacturer of a strobe unit, a conversation with a representative of the FHWA, and conversations with other interested individuals. Each of the jurisdictions on that list was asked if, to its knowledge, any other jurisdiction used SSRIs. Each addition jurisdiction thus identified was then asked the same question. Because of time constraints, the jurisdictions identified in this second round were not contacted.

Data Collection

A cover letter and a questionnaire were mailed to each jurisdiction identified as having used SSRIs (except for those jurisdictions identified in the second round). Data were requested for three accident categories: rear-end accidents, right-angle accidents, and total accidents.

Usable data were not obtained from all of the contacted jurisdictions because

1. Not all jurisdictions responded;
2. Not all of the contacted jurisdictions had experience with SSRIs, despite their identification by others; and
3. Of those jurisdictions that did respond and that did have experience with SSRIs, not all had retrievable accident and volume data.

Usable data were obtained for only 10 intersections, despite the fact that SSRIs have been used at a minimum of 57 locations. It should be noted that it is quite possible that a number of jurisdictions that have used SSRIs were not identified by this project.

DATA ANALYSES

Reasons for Strobe Installation

The responding jurisdictions gave several reasons for installing SSRIs. In general, the reasons fell into two categories: excessive accident experience and the need to draw driver attention to an unexpected signal. Those jurisdictions that cited a specific type of accident as excessive listed right-angle accidents or rear-end accidents as the excessive category.

Statistical Analyses

The usable data were examined by means of statistical tests. The tests used and the results obtained are documented hereafter. It should be noted that the data from the year of SSRI installation were removed from the data base before the performance of statistical tests. This removal was thought necessary because there was no way to increase those data to represent a full year's data without distortion. This was particularly true in those cases where an SSRI had been installed near the beginning or the end of a year.

Sign Tests by Accident Category

The mean number of accidents per year before SSRI installation was compared with the mean number of accidents per year after SSRI installation. It was thought that parametric before and after comparisons for each intersection would have little meaning because each sample had a maximum of three entries. It was also thought that aggregating the data over all intersections would lead to a bias in the data toward intersections with higher accident numbers. For example, a 50 percent reduction in accidents at a location with 50 accidents per year would have been given more statistical weight than a 90 percent reduction in accidents at a location with 10 accidents per year. For these reasons, the sign test was used to compare before and after data.

Strictly speaking, the null hypothesis was that, for each accident category, the after data do not have a different mean value than the before data. Only the "right-angle accidents for the intersection as a whole" category yielded a probability of less than 5 percent, and only one other category (right-angle accidents involving at least one vehicle from an approach with a strobe) yielded a probability of less than 10 percent. Thus, despite the fact that decreases outnumbered increases in five of the six categories, the results are not statistically impressive.

Sign Tests by Accident Rate Category

The mean number of accidents per year before SSRI installation was compared with the mean number of accidents per year after SSRI installation. The null hypothesis for these tests was that the after data for each accident rate category do not have a different mean value than the before data for that category.

For the "approach with strobe" categories, the rates were found by dividing the number of accidents by the average daily traffic (ADT) on all approaches with SSRIs and then multiplying the quantity thus obtained by 1 million. For the "intersection as a whole" categories, the rates were found by dividing the number of accidents by the total approach ADT for the intersection and then multiplying the quantity thus obtained by 1 million.

Despite the fact that decreases led increases in all six categories, only one category (rate of rear-end accidents involving at least one vehicle from an approach with a strobe) yielded a probability of less than 10 percent. Thus, as was the case with the preceding tests, the results are not statistically impressive.

Sign Tests by Intersection

The mean yearly accidents and accident rates before SSRI installation were compared with mean yearly accidents and accident rates after SSRI installation. It was believed that some individual inter-

sections might have experienced a significant drop in accidents and accident rates following SSRI installation, despite the outcome of the previously described sign tests. For this reason, a sign test was performed for each intersection.

Strictly speaking, the null hypothesis being tested was that mean yearly accident experience following SSRI installation is not different from mean yearly accident experience before SSRI installation. Only four categories were used in these tests: rear-end accidents involving at least one vehicle from an approach with a strobe, rate of rear-end accidents involving at least one vehicle from an approach with a strobe, right-angle accidents involving at least one vehicle from an approach with a strobe, and rate of right-angle accidents involving at least one vehicle from an approach with a strobe. Because at least one of these four categories was a subset of each of the other eight categories, it was necessary to eliminate the other eight categories to maintain independence of the observations in each sample.

Decreases led increases at 7 of the 10 intersections. Furthermore, at six of those seven intersections, no one of the four categories showed an increase. Because of the limited number of observations, however, none of the probabilities yielded by the sign test was below 6.25 percent. Perhaps surprisingly, increases led decreases at two of the intersections.

t-Tests

The mean yearly accident rate before SSRI installation was compared with the mean yearly accident rate after SSRI installation. It was hoped that the previously suspected bias in parametric tests toward locations with high accident experience could be reduced by the use of accident rate data. For this reason the data were aggregated over all of the intersections and the mean for each category before SSRI installation was compared with the mean for the same accident category following SSRI installation.

Decreases occurred in five of the six categories, and an increase occurred in the other category. However, none of the differences were significant at the 10 percent level of significance.

Jurisdictional Opinions

Officials of some of the responding jurisdictions thought that the SSRI had been effective; others expressed uncertainty. An official in one jurisdiction stated that the SSRI had seemed to be effective at two intersections but ineffective at a third; in another jurisdiction an official thought that effectiveness was reduced as local drivers became accustomed to the SSRI. In summary, there was no clear consensus on this question.

Other Comments

A summary of other comments made by officials of the responding jurisdictions follows.

1. Maintenance problems with SSRIs had been experienced in three jurisdictions,
2. In three jurisdictions it was thought that SSRI use should be limited to avoid dilution of SSRI effectiveness, and
3. An official of the jurisdiction that used the "vertical slit in center of lens" type of SSRI stated that the jurisdiction had installed SSRIs on all three signal indications on an approach and that a number of drivers had interpreted the indications as flashing red.

SOURCES OF ERROR

There were a number of potential sources of error in this project including the following.

1. Changes in accident reporting criteria may have caused a bias in the data.
2. Volume data were not always provided. When necessary, linear interpolation was performed on volume data. This interpolation was clearly less than desirable, but it was thought to be a reasonable method of increasing the data base. Volume data were never extrapolated.

CONCLUSIONS

In answer to the four questions posed earlier, the following answers are provided.

1. SSRIs have been used in a number of jurisdictions. The confirmed users are the city of Tampa, Florida; Metropolitan Dade County, Florida; the city of Portland, Oregon; Montgomery County, Maryland; the city of Charleston, South Carolina; the state of North Carolina; the state of Maryland; the state of Kansas; the state of New York; and the District of Columbia. There may well be other users who were not identified in this project.
2. SSRIs have been used in attempts to reduce excessive accident rates and to draw driver attention to an unexpected signal.
3. In most cases, the number of accidents and accident rates decreased following SSRI installation. Little statistical significance can be attached to these decreases. However, the lack of statistical significance may be due, at least in part, to the limited amount of data available.
4. Officials of some jurisdictions believe that SSRIs have been effective, but others are uncertain. There is no clear consensus on this question.

ACKNOWLEDGMENTS

The computer time for this project was provided by the Computer Science Center of the University of Maryland. The author wishes to thank all the responding jurisdictions for their cooperation and assistance with this project.

REFERENCE

1. Manual on Uniform Traffic Control Devices for Streets and Highways. National Advisory Committee on Uniform Traffic Control Devices, FHWA, U.S. Department of Transportation, 1978.

Publication of this paper sponsored by Committee on Traffic Control Devices.

Abridgment

Flashing Signal Accident Evaluation

JAMES C. BARBARESSO

ABSTRACT

In this study the relative accident impacts of flashing signal operation and stop-and-go signal operation in Oakland County, Michigan, were evaluated. Analyses were conducted to determine (a) if an accident problem exists at intersections where signals are in a flashing mode during off-peak, nighttime hours; (b) what levels of accident experience can be expected under different conditions and signal operations; and (c) appropriate criteria for making signal-operation decisions for off-peak, nighttime hours. The results of the study indicated that right-angle accidents are significantly overrepresented at four-legged, arterial intersections when signals are in a flashing mode during nighttime hours. A before-and-after analysis demonstrated that the severity of up to 100 percent of late night right-angle accidents can be reduced by eliminating flashing signal operation, with no significant effect on the frequencies of other accident types. The elimination of flashing signal operation appears to be effective in reducing nighttime, right-angle accident frequency and road agency liability exposure both at individual locations and systemwide. Factors that were found to be related to the level of right-angle accidents at flashing signal locations include (a) intersection type (i.e., three-legged or four-legged); (b) the functional classification of the intersecting roadways; (c) the hourly volume ratio (i.e., main street traffic volume to minor street traffic volume); (d) driver impairment; and (e) time of night. When making decisions regarding signal operation during off-peak, nighttime hours, right-angle accident frequency and rate should be primary factors in the decision-making criteria. Other criteria for making signal-operation decisions are presented for use when conditions favor the occurrence of right-angle accidents but none occurred during the review period.

During off-peak, nighttime periods, traffic signals are often placed on flashing operation (i.e., amber in one direction and red in the other) to reduce delay on the major street approaches to intersections. Although this practice improves traffic flow, flashing operation of traffic signals may have detrimental effects on traffic safety under certain conditions.

An analysis was conducted to determine

1. If an accident problem exists at intersections where signals are placed on a flashing mode during off-peak, nighttime hours;
2. What levels of accident experience can be expected under different conditions and signal operations; and
3. Appropriate criteria for the development of signal-operation procedures during off-peak, nighttime hours.

The relative accident impacts of flashing versus stop-and-go (i.e., standard green-amber-red cycle) signal operation in Oakland County, Michigan, were evaluated. The impetus for this study involved both safety and liability questions. The intent of this study was to determine how the Oakland County Road Commission could improve its liability posture and cost-effectively mitigate potential hazards on the county road system.

The study was conducted in two stages. The first stage involved a before-and-after analysis of accidents at six four-legged intersections where the hours of stop-and-go signal operation had been extended or where flashing operation had been eliminated. The next stage was a comparative analysis (i.e., with-and-without study) of accidents at intersections categorized by signal operation (i.e., flashing versus stop-and-go), by intersection type (i.e., three-legged, four-legged) and by the functional classification of the intersecting roadways.

PREVIOUS STUDIES

The FHWA report entitled "A Study of Clearance Intervals, Flashing Operation and Left-Turn Phasing at Traffic Signals," Volume 3 (1) provides the most comprehensive review of the relative accident impacts of flashing and stop-and-go signal operation. An extensive review of previous research is included in that report.

The primary result obtained in the FHWA study was that the number of right-angle accidents was significantly higher at flashing signal locations than at stop-and-go locations. The results of the FHWA study were used to derive guidelines for using flashing signal operation. A critique of these guidelines is included in this paper.

METHODOLOGY

The first stage of this study consisted of a before-and-after study of six signalized, four-legged intersections. The six study sites for the before-and-after analysis were chosen at random from a listing of pretimed signals where flashing operation had been eliminated. The only restrictive criterion in the selection of the study sites was that accident data be available for 3 years before and after the signal operation change. Paired t-tests were performed for the six study sites to determine if accident frequency and accident rate per million vehicles changed significantly in the after period. Accident types were categorized as right-angle accidents, left-turn accidents, rear-end accidents, and other accidents. An additional 10 intersections, where signals remained on flash operation during off-peak, nighttime hours throughout the study period, were randomly chosen to provide a control group for the before-and-after study and to supplement the analysis of other factors that may have some influence on accident levels. These factors included

1. Hourly intersection traffic volume;
2. Main street hourly volume to minor street hourly volume (i.e., the volume ratio) (2); and
3. Drinking involvement.

The second stage of the analysis consisted of a with-and-without study to compare the mean right-angle accident rates and frequencies per year-hour of flashing signal locations and stop-and-go locations. Flashing signal locations were categorized by intersection type and the functional classification of the intersecting roadways (Table 1).

TABLE 1 Sample Sizes for Flashing Signal Location Categories

	Functional Classification		
Intersection Type	Arterial-Arterial	Arterial-Collector	Total
Four-legged, right-angle	29	17	46
Three-legged, T	22	14	36
Total	51	31	82

For each of these intersection types the mean frequency and rate of right-angle accidents per year-hour were calculated for hours when the signals flash. T-tests were conducted to determine if the means differed significantly from each other and from the mean for the hours of 11 p.m. to 6 a.m. at a sample of 21 four-legged intersections where the signals operate on a stop-and-go basis 24 hr a day. Accident data for 3 years were analyzed for all intersections. The statistical tests were conducted at the 0.01 level of significance.

RESULTS

The results of both the before-and-after study and the with-and-without study clearly indicated that significant reductions in nighttime right-angle accident frequency and rate can be attained by eliminating flashing signal operation at four-legged intersections of two arterial roadways. Four-legged intersections of arterial roadways where signals flash during off-peak, nighttime hours experienced significantly greater frequencies and rates of right-angle accidents than other intersection types. Table 2 gives a summary of the results of the with-and-without study. Other results obtained include

1. The rate of right-angle accidents for volume ratios of 2 to 1 or less was significantly higher than the rate for volume ratios of 4 to 1 or greater at flashing signal locations. This result confirms the findings of previous studies.
2. Hourly intersection traffic volume had a negligible impact on right-angle accident frequency during hours of flashing operation.
3. Drinking involvement was significantly overrepresented in right-angle accidents at flashing signal locations.
4. Right-angle accidents at flashing signal locations peaked between midnight and 3 a.m., after which they dropped dramatically. Right-angle accidents at stop-and-go locations peaked between 2 and 3 a.m. (bars close at 2 a.m. in Michigan).
5. Although it was found that rear-end accident frequency was significantly higher at stop-and-go locations during late night hours, no significant difference in rear-end accident rates per million vehicles was found between the two operating modes. Therefore, the difference in rear-end frequencies may be attributable to the relative volumes of traffic at stop-and-go and flasher locations.

The results of this analysis confirm those of earlier studies that demonstrated significant differences in right-angle accident levels between the two operating modes. However, the results of this study indicate that the functional classification of the intersecting roadways and intersection type also influence right-angle accident levels at flashing signal locations.

OPERATIONAL CRITERIA

Although criteria exist for changing signal operation from stop-and-go to flash (3), criteria do not exist for reverting from flashing operation to stop-and-go operation. Because the results of this study indicate that right-angle accident frequency is significantly higher at four-legged, arterial intersections when signals flash during nighttime hours, the hourly frequency of right-angle accidents should be a primary factor in the development of criteria for eliminating flashing signal operation.

The FHWA study cited earlier suggested an accident warrant for eliminating flashing signal operation. However, the suggested warrant was erroneously based on the critical levels of right-angle accident frequency and rate at flashing signal intersections. Because the objective of eliminating flashing signal operation is to reduce right-angle accidents to levels experienced at stop-and-go locations, accident warrants should be based on the critical level of right-angle accidents at stop-and-go locations not flashing signal locations. Otherwise, many flashing signal intersections with right-angle accident problems, relative to stop-and-go intersection standards, would be overlooked for treatment. In summary, the accident warrants suggested by the FHWA are much too restrictive and should be lowered considerably.

Right-angle accident frequency provides a basis for reacting to an accident problem by altering the flash schedule. However, right-angle accidents during flashing signal operation are rare events, and some locations that exhibit conditions favoring

TABLE 2 T-Test Results: Comparative Right-Angle Accident Frequencies

Hourly Mean	Intersection Type (operation)	Arterial-Arterial, 4-Legged (flash)	Arterial-Collector, 4-Legged (flash)	Arterial-Arterial, T (flash)	Arterial-Collector, T (flash)	Arterial-Arterial, 4-Legged (stop-and-go)
0.224	Arterial-arterial, 4-legged (flash)		SIG	SIG	SIG	SIG
0.049	Arterial-collector, 4-legged (flash)			SIG	SIG	NS
0.019	Arterial-arterial, T (flash)				SIG	SIG
0.000	Arterial-collector, T (flash)					SIG
0.092	Arterial-arterial, 4-legged (stop-and-go)					

Note: α = 0.01, SIG = significant difference, NS = not significant.

right-angle accident occurrences may not be accident sites during the review period. The results of this study indicate that a high-risk situation occurs at four-legged intersections of two arterial roadways when traffic signals are in a flashing mode. Therefore, functional classification and intersection configuration provide appropriate surrogate criteria for making signal-operation changes during nighttime periods. (Although they were not analyzed in this study, arterial intersections with more than four legs should also be considered for the elimination of flashing operation.)

Other factors related to right-angle accidents at flashing signal locations include the time of night and the volume ratio. As mentioned earlier, right-angle accidents at flashing signal locations dropped dramatically after 3 a.m. In addition, four-legged intersections with hourly volume ratios of less than 2 to 1 demonstrate significantly higher rates of right-angle accidents than those with ratios greater than 4 to 1 when signals are flashing.

Although sight distance was not analyzed in this study because none of the sample flashing signal locations exhibited sight restrictions from a stopped position, eliminating flashing operation of signals at intersections where sight distance is limited should be considered. Minimum sight distance can be determined using the computational procedures outlined by AASHO (4).

IMPACTS OF FLASH ELIMINATION

The right-angle accident reduction benefits resulting from flash elimination must be weighed against the expected advantages. Increased delay will result, and rear-end accidents, which are generally less severe than right-angle accidents, may increase. However, these disadvantages can be minimized through signal optimization, synchronization, altering cycle lengths, or semiactuation.

Eliminating flashing signal operation will also result in increased emissions of hydrocarbons and carbon monoxide. The total tonnage increases in these pollutants would appear to be significant when analyzed at all intersections in question for a period of 1 year or longer. However, short-term (1 or 8 hr) concentrations should not measurably change, and people will not be affected by an increase in air pollutants.

HUMAN FACTORS

Human factors must be considered when making signal-operation changes during off-peak, nighttime hours. Of primarily importance in this regard are (a) driver impairment, (b) driver expectation, and (c) driver frustration.

Right-angle accidents involving impaired drivers are overrepresented at flashing signal locations relative to stop-and-go locations during the same nighttime hours. This conclusion may indicate a possible perception problem for impaired drivers when faced with a flashing signal. Further research is necessary to determine if this is the case. Regardless, driver impairment must be considered. Signals are normally placed on flashing operation during time periods when drivers are most apt to be tired or under the influence of drugs or alcohol. Stop-and-go operation should be considered until at least 1 hr after bars close.

Another factor to be considered is that of driver expectation. A well-established practice in traffic engineering is to provide drivers with uniform traffic control devices, thereby decreasing driver confusion and enhancing driver expectancy. Flashing signals provide drivers with a set of stimuli that differ from those that they encounter during normal, daytime driving. This may lead to confusion of drivers faced with a flashing signal.

Finally, driver frustration can be expected when drivers are forced to stop for signals during nighttime, low-traffic periods. Such situations are thought to breed contempt and disregard for traffic signals, although documented evidence to support this argument is lacking. However, the evidence suggests that drivers are more apt to stop for a steady red signal, thus reducing the chances of an accident. Nevertheless, attempts should be made to reduce delay and driver frustration through the signal timing alternatives mentioned previously.

LIABILITY IMPLICATIONS

General trends among transportation agencies include the recent surge in litigation and growing concern for reducing liability exposure. Plaintiffs have sometimes argued that flashing signal operation was a casual factor in right-angle accidents at intersections. When this factor is coupled with allegations of limited corner sight distance, the road agency stands to lose a great deal of money.

To reduce road agency liability exposure, it is necessary to identify and treat areas of risk. The analyses in this study indicated that the risk of right-angle accident occurrences at four-legged arterial intersections is higher when signals operate in a flashing mode than when they operate in a stop-and-go mode. Thus, treatment could be justified from a risk management standpoint.

CONCLUSIONS

This study has evaluated the relative accident impacts of flashing signal operation and stop-and-go signal operation during off-peak, nighttime hours. The results of this study indicate that right-angle accidents are significantly overrepresented at four-legged, arterial intersections when signals are in a flashing mode during nighttime hours. T-type intersections and arterial-collector intersections, where signals flash part time, experience significantly fewer right-angle accidents than the other intersection types analyzed.

Other factors that were found to be related to the rate of right-angle accidents at intersections where signals flash at night include the volume ratio, driver impairment, and time of night. Four-legged intersections with hourly volume ratios of 2 to 1 or less (typical for four-legged arterial intersections) experience a significantly greater number of right-angle accidents than those within volume ratios of 4 to 1 or greater. In addition, drinking involvement was overrepresented in right-angle accidents at flashing signal locations relative to stop-and-go locations. Further research is necessary to determine why. Finally, right-angle accident levels at flashing signal locations dropped dramatically after 3 a.m. Consideration should be given to extending stop-and-go operation until at least that time.

Although it was not analyzed in this study, sight distance is an important consideration in the review of intersection signal operation. Eliminating flashing signal operation at intersections with sight restrictions should reduce right-angle accident frequency.

Right-angle accident frequency and rate should be major factors in the development of criteria for

eliminating flashing signal operation. Accident warrants should be based on the critical levels of right-angle accidents at stop-and-go locations, not flashing signal locations. Surrogate criteria may also be used to make decisions regarding nighttime signal operation. The results of this study indicate that a high-risk situation occurs at four-legged intersections of two arterial roadways when signals are in a flashing mode. Therefore, appropriate criteria could include both intersection type and the functional classification of the intersecting roadways.

The elimination of flashing signal operation at high-risk locations during off-peak, nighttime hours appears to be effective in reducing right-angle accident frequency and road agency liability exposure at individual locations or systemwide. Of course, trade-offs must be made in terms of vehicle delay, potential increases in rear-end accidents, and possibly in terms of driver disregard for signals. However, these disadvantages can be effectively dealt with by other means.

Discussion

Olga J. Pendleton*

The primary purpose of this discussion is to recommend a statistical method, which is more powerful in detecting true differences in accident frequencies in before-and-after designs, to be used on the type of data collected in this study.

The statistical methods that were used in this study are quite common and widely practiced in the accident frequency analysis of before-and-after designs. These methods, unfortunately, require stringent normal distribution assumptions or large sample sizes to effectively determine differences in accident frequencies. These statistics, namely the t-test for comparison of means and the z-test for comparison of proportions, are widely used because they are easy to apply and most researchers encountered them in their training in classical statistics. Unfortunately, variables encountered in the field of accident analysis (i.e., accident frequencies and rates) are not normally distributed and the sample sizes required to appeal to asymptotic normal theory are not practically attainable. The assumptions in performing a t-test include, for example, the assumption that the variances of the two comparison populations are equal. According to the data presented in Tables 1 and 2, for example, this assumption appears to be violated in this study. The sample variances of the before-and-after sites differ in magnitude by factors of 272 and 479, respectively, and far exceed the critical value of 4, which is the maximum amount these variances can differ to validate the homogeneity assumption.

The method of analysis proposed in this discussion is the analysis of a before-and-after design with a comparison group, and the recommended statistical procedure for estimating treatment effectiveness is the cross-product ratio. The details of this method can be found in a publication by Griffin (5), and the advantages of this design along with other statistical tests for evaluating treatment effectiveness can be found in the Accident Research Manual (6).

*Texas Transportation Institute, Texas A&M University, College Station, Tex. 77843.

The before-and-after design with a comparison group could be implemented in this study as follows: Suppose the traffic signal at a particular site was in flash mode from 11 p.m. to 6 a.m. before June 1 and then returned to normal operation after June 1. A comparison group would be defined as the number of accidents that occurred between 6 a.m. and 11 p.m. before and after June 1. Then one could address the question: Was there a significant change in accident frequency, caused by the flashing mode operation, above and beyond any change in accident frequency that might have been expected at that site because of normal fluctuations in accident patterns? That is, if the number of accidents that occurred between 6 a.m. and 11 p.m. increased after June 1, the number of 11 p.m. to 6 a.m. accidents might have been expected to increase during this time also. If so, was this increase of the same order of magnitude or did the signal change from flashing to stop and go result in a significant decrease relative to the comparison group? If the 11 p.m. to 6 a.m. accident frequency decreased over this period (Figure 1), the effect of the signal operation mode change would be even more dramatic.

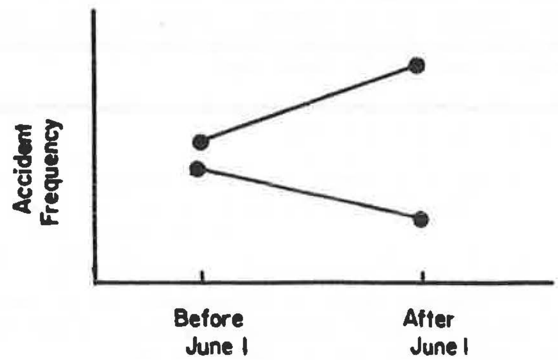

FIGURE 1 Changes in accident frequency.

A hypothetical numerical example will be used to illustrate this method. Consider the following accident frequency table that might reflect the relationship shown in the figure.

	No. of Accidents	
	6 a.m. to 11 p.m.	11 p.m. to 6 a.m.
Before June 1	20	15
After June 1	30	5

The number of 11 p.m. to 6 a.m. accidents was reduced by 67 percent after the signal was returned to stop-and-go mode. However, this reduction underestimates the effect of signal change because it does not take into account the fact that there was a 50 percent increase in 6 a.m. to 11 p.m. accidents during this same period (probably because of increased traffic volume). A better estimate of the percent reduction is obtained using the cross-product or log-odds ratio, which is the ratio of the 6 a.m. to 11 p.m. accidents to the 11 p.m. to 6 a.m. accidents. That is,

$$CPR = (20/30)/(15/5) = [(20)(5)]/[(30)(15)] = 0.22 \quad (1)$$

This reduces to a ratio of the product of the diagonals in the table. The estimate of the percent reduction is then

$$(CPR - 1) \times 100 = 78 \text{ percent} \quad (2)$$

or there was a 78 percent reduction in the number of

accidents between 11 p.m. and 6 a.m. after June 1. If (CPR - 1) had been positive, this would have meant there was an increase in 11 p.m. to 6 a.m. accidents.

Obtaining the estimate is only part of the analysis. The next step is to ask if this reduction or increase was statistically significant. This is equivalent to asking if the CPR is equal to 1. The test is a normal z-test and is computed as

$$z = \ln(\text{CPR})/[(1/20) + (1/30) + (1/15) + (1/5)]^{1/2} = -2.54 \quad (3)$$

which is significant at the 5 percent level.

This discussion has focused on the analysis of a single site; methods exist for combining sites, but such methods should be used with caution.

This study attempted to use a control group defined as a similar site over the same time period. The selection of a comparison group defined by complementary hours at the same site has the advantage that differences in exposure or roadway geometrics are minimal because the site of the comparison group is the same. Of course, inherent in this is the assumption that these site variables are equal for the two daily time periods. This is sometimes a problem, especially for the variable exposure. However, most exposure variables, such as average daily traffic or vehicle-miles of travel, are only available in daily units, not hourly, so they are difficult to control for in any analysis. They surely are not reflected in the simple t-test on means or the z-test on proportions using the before-and-after experience of the test site alone.

In summation, the before-and-after design with a comparison group could be easily implemented in this study and is more powerful in evaluating the effect of changing flashing signals to stop and go. Another powerful design would require changing these same signals back to flashing and examining a before-and-after design with a comparison group.

Author's Closure

Pendleton's discussion is an informative dissertation on research design and statistical methods for accident countermeasure evaluations. As indicated by Pendleton, more powerful research designs than those used in the evaluation are available to researchers. Nevertheless, it is unlikely that the results would have been drastically different if the designs or statistical methods suggested by Pendleton had been used in this study.

Pendleton is correct in that the use of a t-test in before-and-after designs requires a normal distribution assumption. As a general rule, accident frequencies more closely follow a Poisson distribution. However, applying a Poisson comparison of means test will demonstrate similar, significant results.

Pendleton also discusses the advantages of the before-and-after design with a comparison group. A variation of this research design was used in this study. The Accident Research Manual (6) describes a comparison group as a culled sample of locations demonstrating qualities similar to those of the test group locations before treatment. In this study, the test group was randomly chosen from a listing of intersections where flashing signal operation was eliminated in the after period. The comparison group was randomly chosen from a listing of intersections that retained flashing signal operation in the after period. Admittedly, the two groups were not identical, but they were very close to equal before treatment of the test group.

In conclusion, it is unlikely that different study designs or statistical methods would have resulted in less significant results. In fact, two different research designs and a variety of statistical methods were used in this study, and the same conclusion was derived from both study designs. It was clear that the elimination of flashing signal operation is effective in reducing late night, right-angle accidents at intersections.

The reader need only review the data to observe the dramatic effects of eliminating flashing signal operation. For example, in two-thirds of the before-and-after test sites, 100 percent reductions in right-angle accidents were noted after flashing signal operation was eliminated. The average accident reduction for all test sites was greater than 97 percent. These reductions were experienced even though right-angle accident frequency at the comparison sites increased by almost 36 percent during the study period.

At one of the comparison sites, an average of more than 10 right-angle accidents occurred annually during flashing signal operation. Flashing signal operation was eliminated at this location in 1982. In the year following flash elimination, no right-angle accidents occurred during the hours in which the signal had operated on flash. Total accidents at this intersection (for a 24-hr period) were reduced by more than 33 percent. Given these and similar examples of dramatic reductions in accident frequency, statistical analyses may not even be necessary to demonstrate the effectiveness of eliminating flashing signal operation.

REFERENCES

1. B. Benioff, C. Carson, and F.C. Dock. A Study of Clearance Intervals, Flashing Operation and Left-Turn Phasing at Traffic Signals, Volume 3. FHWA-RD-78-48. FHWA, U.S. Department of Transportation, Jan. 1978.
2. J.A. Marson. A Comparison of Accident Experience at Intersections with Flashing and Regular Traffic Signal Control During Low-Volume Time Periods. M.S. thesis. Michigan State University, East Lansing, 1976.
3. Michigan Manual of Uniform Traffic Control Devices. Michigan Department of Transportation and the Michigan Department of State Police, Lansing, 1981.
4. A Policy on Geometric Design of Rural Highways. AASHO, Washington, D.C., 1965.
5. L.I. Griffin. Three Procedures for Evaluating Highway Safety Improvement Programs. TARE 47. Traffic Accident Research and Evaluation Program, Texas Transportation Institute, Texas A&M University, College Station, 1982.
6. F.M. Council, D.W. Reinfurt, B.J. Campbell, F.L. Roediger, C.L. Carroll, A.K. Dutt, and J.R. Dunham. Accident Research Manual. FHWA/RD-80/016. FHWA, U.S. Department of Transportation, 1980.

The opinions and viewpoints expressed in this paper are entirely those of the author. They do not necessarily reflect the viewpoints or policies of the Oakland County Road Commission or any federal or state transportation agencies.

Publication of this paper sponsored by Committee on Traffic Control Devices.

Abridgment

Analysis of Delay and User Costs at Unwarranted Four-Way Stop Sign Controlled Intersections

MICHAEL N. BYRD and DONALD B. STAFFORD

ABSTRACT

A study was conducted to examine the operational characteristics of traffic controls at low-volume, low-speed intersections with unwarranted four-way stop sign control. The research involved the collection of traffic volume and delay data at eight selected four-way stop sign controlled intersections in three cities in northwestern South Carolina. These data and various unit delay values and unit cost factors were used to quantify the delays and road user costs experienced by motorists at these intersections. The delay and road user costs that would result from two-way stop sign control at the intersections were estimated based on field data collected at similar intersections and compared with the delay and road user costs associated with four-way stop sign control. This comparison of delay and road user costs showed that four-way stop sign control at the eight intersections selected for study caused 26,430 hr of additional delay and $296,610 in additional road user costs annually. Four-way stop sign control produced an average of 3,300 hr of additional delay per intersection and $37,080 per intersection in additional road user costs annually. Therefore, it was concluded that unless an accident problem susceptible to correction by four-way stop sign control exists, the unwarranted use of four-way stop sign control results in unnecessary delay and road user costs to the driving public and that the intersection traffic control should be changed to two-way sign control. Also, traffic engineers should resist efforts by the public to have four-way stop sign control installed at intersections where it is not warranted.

The purpose of the stop sign is to inform the motorist of a requirement to stop. Ideally, stop signs should protect the continuous flow of through traffic on the major route. Stop signs should not be installed for speed control (1). Nevertheless, community groups often exert pressure on their elected officials to install stop signs, particularly unwarranted four-way stop signs, for speed control. Citizens and public officials believe that four-way stop sign control will reduce vehicle speeds and significantly increase the safety of their community, but they apparently are unaware of the magnitude of the increased delay and additional road user costs associated with this type of intersection traffic control. Although the stop sign is not recommended for speed control, some traffic engineers apparently approve four-way stop sign installations because of a lack of concern for the increased delay and higher road user cost involved or because of their desire to demonstrate a concern for community safety by installing a traffic control device that has a reasonably low initial cost. However, it should be noted that under some circumstances warranted four-way stop sign control of intersections can provide accident savings that justify the increased delay and higher road user cost.

Unwarranted four-way stop sign control causes greater delay and road user costs than does two-way stop sign control because of the requirement that all vehicles come to a complete stop. In addition, increased fuel consumption and vehicle emissions result from the unnecessary stopping maneuvers associated with unwarranted four-way stop sign control. Therefore, the overall impact of unwarranted four-way stop sign control is undesirable from several standpoints. The installation of four-way stop sign control should be strictly limited to intersections that clearly warrant this type of traffic control.

PURPOSE AND OBJECTIVES

The purpose of the study reported here was to evaluate certain operational characteristics of traffic at four-way stop sign controlled intersections that did not meet accepted minimum warrants for the installation of four-way stop sign control. To fulfill this purpose, the specific objectives were (a) to collect traffic volume and delay data at selected four-way stop sign and two-way stop sign controlled intersections, (b) to quantify delay and road user costs at the selected four-way stop sign controlled intersections, (c) to estimate the delay and road user costs associated with two-way stop sign control at the selected four-way stop sign controlled intersections, and (d) to present information on the comparison of delay and road user costs at intersections with different forms of intersection traffic control.

DATA COLLECTION

The primary criterion for selecting the four-way stop sign controlled intersections for this study was the failure of an intersection to meet the minimum warrants for multiway stop sign control (1). The eight four-way stop sign controlled intersections selected for study were located in Anderson and Oconee counties in northwestern South Carolina. Three intersections were located in the Oconee County town of Seneca that had a 1980 population of 7,436. Three intersections were located in Walhalla, the county seat of Oconee County, that had a 1980 population of 3,977. Two intersections were located inside the corporate limits of Anderson, a city that had a 1980 population of 27,965.

Data were collected on traffic volume and vehicle

delay characteristics at each intersection during weekdays so the data would represent typical traffic conditions. The two components of intersection delay measured were stopped-time delay and the delay associated with speed change cycles. The procedures followed in conducting the stopped-time delay studies were those described in the Manual of Traffic Engineering Studies (2); these procedures represent a widely adopted method of obtaining information on stopped-time delay. This method involved the counting of vehicles stopped on an intersection approach at successive time intervals.

Speed change cycle delay, which is defined as the additional time required to decelerate from a specific approach speed and then accelerate back to the initial approach speed above the time required to travel through the intersection at the initial approach speed, was computed for each intersection approach. In this analysis all vehicles approaching a four-way stop sign controlled intersection were assumed to be traveling at the speed limit and then come to a complete stop. The unit values of delay used in the analysis of the speed change cycle delay were obtained from Winfrey's Economic Analysis for Highways (3). The speed change cycle delay was combined with the stopped-time delay to obtain total intersection delays.

The operational restrictions imposed by intersection traffic control devices create additional road user costs for motorists. It has long been recognized that traffic congestion and the associated delay are not only inconvenient but add to the cost of transporting goods and passengers. A highway improvement that reduces congestion and delay can result in a significant reduction in road user costs. Therefore, quantifying road user costs is an important part of a study of the economics of intersection traffic control. The two types of costs considered in this study were those associated with the operation of the vehicle and the value of motorists' time. Vehicle operating costs are defined as those direct road user costs that result from the operation of a vehicle. Because all vehicles must stop at four-way stop sign controlled intersections, the vehicle operating costs calculated were those resulting from vehicles decelerating from the initial speed to a stop and accelerating back to the initial speed plus vehicle idling costs while stopped. This study used the widely accepted methods provided by the AASHTO "Red Book" (4) for quantifying vehicle operating costs.

The AASHTO methods involved the application of cost factors to the traffic volume and delay data collected at each four-way stop sign controlled intersection. AASHTO provided cost factors that represented 1975 conditions and price levels. Before applying these cost factors, they were updated to March 1982 conditions and price levels using procedures recommended in the Red Book (4). These procedures are based on the consumer and wholesale price indexes published by the U.S. Bureau of Labor Statistics.

The daily volume of vehicles passing through each intersection was multiplied by the appropriate updated AASHTO cost factor to obtain the speed change cycle costs generated at an intersection. Because of the low percentages of single-unit and tractor-semitrailer trucks in the traffic stream at the intersections involved in the study, the daily speed change cycle costs at the intersections resulted primarily from passenger vehicle speed change cycles.

Idling costs are incurred when a vehicle is stopped at an intersection. The amount of cost is dependent on the vehicle type and the length of time the vehicle is stopped. To calculate the daily vehicle idling cost at an intersection, the unit values of idling cost recommended in the Red Book (4) were updated and used. Each unit cost value per idling hour was multiplied by the daily stopped-time delay to obtain the daily idling costs at an intersection.

This study also placed a value on the time associated with traveling through an intersection by multiplying unit values of time by the amount of time consumed and by a vehicle occupancy factor. The unit values of time used were $3.00 per traveler hour for the occupants of passenger vehicles and $6.00 per traveler hour for the occupants of single-unit and combination tractor-semitrailer trucks. All passenger vehicles were considered to have an average occupancy of 1.6 adults per vehicle and all trucks were considered to have an average occupancy of 1.2 adults per vehicle. The total road user costs at each four-way stop sign controlled intersection were calculated by combining the vehicle operating cost from speed change cycles and vehicle idling with the time cost.

IMPACT OF TWO-WAY STOP SIGN CONTROL

The delay and road user costs associated with four-way stop sign control were compared with the delay and road user costs resulting from two-way stop sign control. Estimation of the delay and road user costs that would result at the four-way stop sign controlled intersections if they were controlled by two-way stop signs required the collection of additional delay data at two-way stop sign controlled intersections. The additional delay data were analyzed to obtain numerical values of the stopped-time delay per vehicle and the frequency and magnitude of all speed change cycles associated with two-way stop sign control. The selection of intersections where the additional delay data were collected was done on the basis of similarity of traffic volumes, approach speeds, turning movements, traffic stream characteristics, and abutting land uses at the intersection and at a particular four-way stop sign controlled intersection examined previously. Each intersection was located as close as possible to a particular four-way stop sign controlled intersection to take advantage of similarities in traffic characteristics and environmental conditions in a given area. Four two-way stop sign controlled intersections were analyzed. Each of the four-way stop sign controlled intersections had a highway in common with one of these two-way stop sign controlled intersections.

The stopped-time delay that was measured at the two-way stop sign controlled intersections was used to estimate the stopped-time delay that would result if the four-way stop sign controlled intersections were controlled by two-way stop signs. The stopped-time delay per vehicle on the minor approaches was determined for each two-way stop sign controlled intersection and these values were applied to the traffic volume on the minor approaches of the four-way stop sign controlled intersections.

The frequencies and magnitudes of the speed change cycles that occurred at the two-way stop sign controlled intersections were applied to the traffic volumes on the four-way stop sign controlled intersections to estimate the speed change cycle delay that would result if the four-way stop sign controlled intersections were controlled by two-way stop signs.

The analysis of the economic impact of two-way stop sign control at the four-way stop sign controlled intersections focused on road user costs. These road user costs consisted of the vehicle operating costs from speed change cycles and vehicle

idling plus the time cost to vehicle occupants. The methods and procedures used to calculate the road user cost that would result from two-way stop sign control were the same as those described previously for four-way stop sign control. A more detailed description of the study techniques employed, the methods used for computing total intersection delay and road user costs, and the results obtained from the study has been presented by Byrd (5).

SUMMARY AND CONCLUSIONS

This study investigated the operational characteristics of selected unwarranted four-way stop sign controlled intersections. Field data were collected on traffic volume and delay characteristics at eight intersections currently operating with four-way stop sign control in northwestern South Carolina. Stopped-time delay studies were conducted on all of the approaches to the intersections. The traffic volume and delay data were then used in conjunction with unit delay and cost factors obtained from previous research to quantify delay and road user costs at each intersection.

The annual delay at the eight four-way stop sign controlled intersections was calculated to be 42,660 hr or 5,330 hr per intersection. Sixty-eight percent of this delay resulted from speed change cycles and the remaining 32 percent was stopped-time delay. The annual road user cost that resulted at all eight four-way stop sign controlled intersections was $477,960 or $59,750 per intersection. Fifty-four percent of this cost resulted from vehicle operating costs during the speed change cycles. Time costs represented 44 percent and vehicle idling costs accounted for the remaining 2 percent of the road user costs.

To estimate the magnitude of the possible benefits of two-way stop sign control at the eight four-way stop sign controlled intersections that did not meet the minimum traffic volume warrant of the Manual on Uniform Traffic Control Devices (1), intersections with comparable traffic characteristics and currently operating under two-way stop sign control were analyzed. Using the data obtained at these intersections, the delay and road user costs that would result from two-way stop sign control at the four-way stop sign controlled intersections were estimated. These estimates were then compared with the delays and road user costs that existed with four-way stop sign control to determine the impact that two-way stop sign control would have on the delay and road user costs at these intersections. These comparative analyses showed that the installation of two-way stop sign control would reduce intersection delay and intersection road user costs 62 percent.

Unwarranted four-way stop sign control at the eight intersections selected for study, assuming 100 percent driver obedience of stop signs, caused 26,430 hr of additional delay and $296,610 in additional road user costs annually. Neglecting differences in traffic volumes at the intersections, unwarranted four-way stop sign control caused an average of 3,300 hr of additional delay per intersection and $37,080 per intersection in additional road user costs annually.

This study was conducted on low-volume, low-speed intersections with unwarranted four-way stop sign control in northwestern South Carolina. In the context of this study, low-volume intersections were considered to be intersections with fewer than 8,000 vehicles per day on all approaches. Low-speed intersections were considered to be intersections with approach speed limits of 35 mph or less. The following conclusions, which were derived from this study, may apply to intersections similar to the ones investigated but not to all intersections with four-way stop sign control.

1. The unwarranted use of four-way stop sign control causes motorists to experience substantial amounts of additional and unnecessary delay and road user costs.
2. The installation of two-way stop sign control at an intersection operating with unwarranted four-way stop sign control can produce significant delay reductions and road user cost savings.
3. The total delay to the minor highway traffic at unwarranted four-way stop sign controlled intersections would not be significantly changed by the installation of two-way stop sign control.

The effectiveness of four-way stop sign control as a safety measure is dependent on individual intersection characteristics. If an intersection is operating with unwarranted four-way stop sign control that is not a needed safety measure, the driving public is forced to pay unnecessary road user costs and suffer delay. In addition, it has been established previously that the use of unwarranted four-way stop sign control results in increased fuel consumption and vehicle emissions. Therefore, this study concluded that unwarranted four-way stop sign control at low-volume, low-speed intersections should be changed to two-way stop sign control when highway safety will not be seriously compromised. Also, traffic engineers should use the information on additional delay and road user costs associated with four-way stop sign controlled intersections to resist the efforts of the public to have four-way stop sign control installed at intersections where it is not warranted.

REFERENCES

1. Manual on Uniform Traffic Control Devices for Streets and Highways. FHWA, U.S. Department of Transportation, 1978.
2. P.C. Box and J.C. Oppenlander. Manual of Traffic Engineering Studies, 4th ed. Institute of Transportation Engineers, Washington, D.C., 1976.
3. R. Winfrey. Economic Analysis for Highways. International Textbook Company, Scranton, Pa., 1968.
4. A Manual on User Benefits Analysis of Highway and Bus-Transit Improvements. AASHTO, Washington, D.C., 1977.
5. M.N. Byrd. An Analysis of Vehicle Delay and Road User Costs at Unwarranted Four-Way Stop Sign Controlled Intersections. M.S. special project report. Clemson University, Clemson, S.C., 1983.

Publication of this paper sponsored by Committee on Traffic Control Devices.

Abridgment

Passing Requirements for Two-Lane Highways in Mountainous Areas

NICHOLAS J. GARBER and MITSURU SAITO

ABSTRACT

The Virginia Department of Highways and Transportation uses a special centerline marking designated mountain pavement marking on two-lane highways in mountainous regions. This marking consists of a single broken line supplemented by PASS WITH CAUTION signs. This marking has been criticized because it does not prohibit passing on sections of highways with inadequate sight distances. Consequently, a study was conducted to determine if this marking system should be replaced with the Manual on Uniform Traffic Control Devices (MUTCD) standard marking pattern and to develop guidelines for minimum lengths of passing zones and minimum sight distances for safe passes on two-lane mountain roads. In this paper the results of that part of the study that dealt with the development of minimum passing criteria are presented. Passing maneuvers were recorded at five sites with a 16-mm movie camera. Relevant data were then extracted from the film and used to develop a regression model for the minimum length of passing zone based on the passing speed and the difference in speeds of the passing and impeding vehicles. A minimum passing sight distance for a safe pass or a comfortable aborting of the pass was then developed using the concepts of critical position and comfortable deceleration. The results indicate that the minimum values suggested in the MUTCD are inadequate and that there are no significant differences in traffic characteristics between roads with the special marking and those with standard MUTCD marking. The results also indicate that, although speed is the major factor affecting the passing distance on two-lane highways in mountainous regions, other factors such as the difference between the speeds of the passing and impeding vehicles and grade, in that order, also have some effect.

The long-standing policy of the Virginia Department of Highways and Transportation is to use the centerline marking standards outlined in the Manual on Uniform Traffic Control Devices (MUTCD) (1). The department, however, uses a special type of marking on two-lane highways in mountainous regions. This special marking, designated mountain pavement marking, consists of a single, broken, yellow line supplemented by PASS WITH CAUTION signs (Figure 1). Passing maneuvers are not prohibited by the solid yellow line, even when sight distances are inadequate for prevailing speeds, so the decision to pass is left entirely to the motorist. The argument in favor of this marking pattern is that it allows motorists to legally pass slow-moving vehicles where it would not be possible to do so for long distances if these roads were marked in compliance with the MUTCD standards.

FIGURE 1 Mountain pavement marking.

In response to criticism of this practice of marking two-lane highways in mountainous regions, a study was undertaken to evaluate the marking and to determine minimum lengths of passing zones and sight distances for these highways.

In this paper the minimum criteria developed for passing zones and sight distances are documented in terms of

1. A description of the regression model developed for minimum passing sight distances and minimum lengths of passing zones;
2. Recommended guidelines for establishing passing and no-passing zones; and
3. Results of a comparison of the before and after data, with the focus on the driver's interpretation of and compliance with the MUTCD pavement marking pattern and the evaluation of the proposed passing zone lengths.

PURPOSE AND OBJECTIVES

The primary purpose of this portion of the study was to develop guidelines that can be used to delineate passing and no-passing zones on two-lane highways in mountainous regions using the MUTCD standard pavement marking pattern. The objectives were

1. To determine safe and acceptable passing distances for this type of roadway,
2. To determine minimum sight distances for safe passing, and
3. To recommend means of determining appropriate marking patterns for the roads.

STUDY APPROACH

The research was designed as a before-and-after study. The before phase investigated traffic operational characteristics and passing maneuvers at selected sites striped with the mountain pavement marking. The data taken at these sites were subsequently used in developing pass models that, in turn, were used to formulate guidelines for minimum passing zone lengths and minimum passing sight distances. Two of the sites used in the before phase and one other site were then selected for the after phase. The mountain pavement marking at these three sites was replaced with the standard MUTCD marking patterns for passing and no-passing zones in conformity with the guidelines developed. During the after phase, motorists' interpretation of the MUTCD marking patterns and compliance with them, and the adequacy of the proposed minimum passing zone lengths, were evaluated. Data on traffic operational characteristics were also collected during this phase.

From an inventory of the roadways striped with the mountain pavement marking, five sites were selected for study based on the criterion that traffic volume, operating speed, and passing maneuvers at the selected sites be representative of the range that exists on mountainous roads striped with mountain pavement markings.

DATA COLLECTION

The data collection consisted of two major tasks: the collection of traffic flow data with an electronic traffic data acquisition system and the filming of passing maneuvers with a 16-mm movie camera.

A traffic data recorder (TDR) system developed by Leupold & Stevens, Inc., was used for the first task. Traffic operational data such as volume, vehicle speeds, headways, traffic queues, and vehicle classifications were collected during a period of at least 24 continuous hours on Tuesday through Thursday.

For the second task, a Canon Scoopic 16 MS 16-mm movie camera was used. The camera was placed at a point from which the centerline pavement marking was clearly visible. A film speed of 24 frames per second was used throughout the study. Kodak Ektachrome film 7241EE (ASA 80), on 100-ft rolls was used throughout the study. Data on passing maneuvers were collected from a total of 85 passing maneuvers filmed.

Figure 2 is a schematic presentation of the passing maneuver. Among the distance elements shown in the figure, PD, D_3, D_9, G_2, and X' were extracted from the passes filmed with the 16-mm camera. The distance elements are defined as

P = passing vehicle;
I = impeding vehicle;
O = oncoming vehicle;
PD = passing distance denoting the distance traveled by the passing vehicle while it is on the left lane;
D_3 = distance traveled by the passing vehicle from the head and tail position, where the passing vehicle catches up with the impeding vehicle, to completion of the pass;
D_9 = distance traveled by the passing vehicle from the abreast position to the completion of the pass;
X' = space headway retained by the passing vehicle just before it encroaches onto the left lane, before spacing;
G_2 = space headway left for the impeding vehicle by the passing vehicle when it completes the pass, after spacing;
C = clearance distance between the passing and oncoming vehicle at completion of passing maneuver;
D_{cp} = distance traveled by the passing vehicle between the critical position and the position of completion of a pass; and
D_4 = distance traveled by the oncoming vehicle while the passing vehicle travels from the critical position to completion of pass.

DATA REDUCTION AND ANALYSIS

Traffic Operating Characteristics

Operating Speed

Eighty-fifth percentile speeds observed were between approximately 40 and 50 mph, and mean speeds were between 35 and 45 mph. The difference in speeds between opposing lanes was found to be significant at 95 percent confidence for grades greater than 5 percent. The maximum speed difference observed between the opposing lanes was 7.0 mph.

Speed Difference Between Passing and Impeding Vehicles

The speeds of impeding vehicles varied from 15 to 45 mph, whereas those of passing vehicles ranged from 30 to 64 mph.

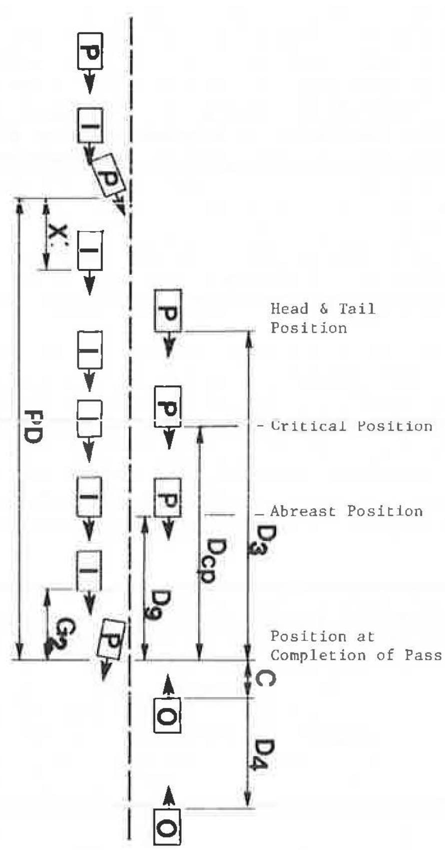

FIGURE 2 Distance elements of passing maneuvers.

Passing Speed Versus Off-Peak 85th Percentile Speed

The MUTCD employs the prevailing off-peak 85th percentile speed as the independent variable to compute minimum passing sight distances. The accuracy of this for the roads in this study was checked by comparing the mean speed of the passing vehicles with the off-peak 85th percentile speed at the appropriate sites, and it was found that in all cases the mean passing speeds were approximately equal to the off-peak 85th percentile speeds.

Regression Analysis

A stepwise multiple linear regression analysis was performed using the software package BMDP (2) with the passing distance as the dependent variable and passing speed, available sight distance, speed difference between passing and impeding vehicles, and grade as independent variables. The analysis showed that passing speed had the greatest impact on passing distance. Passing speed was followed by available sight distance, then speed difference, with grade having the least impact. This analysis, however, also showed that multicollinearity exists between sight distance and speed (i.e., speed is related to sight distance). Further analysis also showed that, for speeds less than 50 mph and grades less than 10 percent, the effect of grade on passing distance is minimal. Therefore, the regression equation was developed using the two major variables, speed and speed difference. The equation thus obtained is

$$PD = 266.397 + 9.689\,V - 12.448\,m \qquad (1)$$

where

PD = passing distance in feet,
V = passing speed (off-peak 85th percentile speed) in mph, and
m = speed difference in mph.

Minimum Lengths of Passing Zones

In this study a passing zone is analogous to the passing distance. It is the distance traveled by the passing vehicle on the left lane in a passing maneuver. Within this distance, the passing vehicle encroaches onto the left lane, passes the impeding vehicle, and returns to the right lane on completion of the passing maneuver.

In developing the proposed minimum lengths of passing zones, two factors were taken into consideration. First, the speed difference (m) used in the regression model was 12 mph. This assures that the lengths of passing zones suggested will be equal to or greater than the actual passing distances of 85 percent of all passing maneuvers. Second, to provide for passing maneuvers that do not commence at the beginning of the passing zone, the 95 percent confidence level upper limit of the obtained regression model was used. The proposed minimum lengths of passing zones for different 85th percentile speeds were, therefore, obtained from this model and are given in Table 1. It can be seen that the proposed minimum lengths are greater than 400 ft, which is the minimum suggested in the MUTCD for all speeds. This suggests that the minimum length of 400 ft suggested in the MUTCD may be inadequate for two-lane, two-way highways in mountainous regions, even at the low speed of 30 mph.

TABLE 1 Suggested Minimum Lengths of Passing Zones for Two-Lane Highways in Mountainous Regions

85th Percentile Speed (mph)	Proposed (ft)	MUTCD (ft)
30	560	400
35	610	400
40	660	400
45	710	400
50	750	400

PASSING SIGHT DISTANCE REQUIREMENTS

In this study, the passing sight distance was defined as the sum of the distance between the critical position (CP) and the position at completion of a pass, the clearance distance between the passing and oncoming vehicles at completion of a pass (C), and the distance traveled by an oncoming vehicle while the passing vehicle travels from the critical position and completely returns to the right lane (D_4). This relation is shown in Figure 2.

To compute the passing sight distances for different passing speeds it was first necessary to locate the critical position of a passing maneuver. The critical position was defined by Lieberman as the point where "the decision by the passing vehicle to complete the pass will afford it the same clearance relative to an oncoming vehicle, as will the decision to abort the pass" (3). As the deceleration rate decreases, the passing motorist must decide earlier to abort.

In order to determine the critical position in this study, the deceleration rates (d) necessary for an abort maneuver starting at a point with a clearance distance (C') equal to the clearance distance (C) for a pass maneuver starting at the same point were calculated for different passing speeds.

The deceleration rates computed for the different positions were then compared with acceptable deceleration rates. It was determined that acceleration rates for the critical position located at 2/3 PD are within acceptable limits, whereas those for critical positions located at less than 2/3 PD tend to exceed comfortable limits when passing speeds are greater than 40 mph. It was, therefore, decided to select 2/3 PD as the critical position, and this position was used to determine the sight distance requirements given in Table 2. The suggested minimum passing sight distances given in Table 2 were computed for equal upgrade and downgrade speeds by the equation

$$PSD = (4/3\ PD) + C$$

where

PD = passing distance, in feet, from regression model m = 12, and
C = clearance distance estimated from AASHTO.

and for different speeds between upgrade and downgrade by the equation

$$PSD' = (2/3\ PD) + C_h + D_h$$

where

PSD' = adjusted passing sight distance, in feet,
C_h = clearance distance for higher speed, in feet, and
D_h = distance traveled by oncoming vehicle at

TABLE 2 Comparison of Minimum Passing Sight Distance Requirements

Upgrade 85th Percentile Speed (V_l) (mph)	Minimum Passing Sight Distance (ft): Equal Upgrade and Downgrade Speeds — This Study	MUTCD	Different Speeds Between Upgrade and Downgrade (this study): $(V_l < V_h < (V_l + 5.0)$	$(V_l + 5.0) < V_h < (V_l + 10.0)$
30	645	500	700	800
35	735	550[a]	800	870
40	825	600	885	950
45	910	700[a]	970	1,070
50	1,000	800	1,095	1,190
55	1,115	900[a]	1,200	—[b]

[a] Interpolated.

[b] Passing sight distances were computed only for passing speeds up to 55 mph because grade may significantly affect passing distance when passing speeds are higher than 55 mph.

higher velocity during time passing vehicle travels from critical point to completion of pass.

A comparison of the minimum sight distances obtained from this study with the corresponding values suggested in the MUTCD indicates that the MUTCD values are inadequate for two-lane, two-way highways in mountainous regions.

EVALUATION OF PROPOSED GUIDELINES

The adequacy of the proposed guidelines when used to provide passing zones marked with the MUTCD standard patterns was evaluated by conducting an after study at two of the five sites selected for the before study and a new site. Only two of the original five sites were used for the after study because only at these sites do the traffic and geometric characteristics conform to the proposed guidelines.

The results indicate that when passing zones were provided based on the guidelines developed, 80 percent of passing motorists returned to the right lane without intruding into the passing zone of the opposing lane. This indicates that the majority of motorists correctly interpreted the MUTCD passing and no-passing zone marking patterns.

At sections where the mountain pavement marking was replaced with the double solid yellow line for the no-passing zone, it was found from the data taken with the electronic data acquisition system that very few passing maneuvers occurred. This indicates that motorists correctly interpreted the marking and complied with it.

The results also indicate that at passing zones based on the guidelines developed, a minimum of 88 percent of the passing maneuvers at each site were completed within the proposed minimum length of passing zone for the 85th percentile speed at the site. These figures suggest that the proposed minimum lengths for passing zones are adequate.

CONCLUSIONS

The results of this study indicate that although speed is the major factor affecting the passing distance on two-lane, two-way highways in mountainous regions, other factors such as the speed difference between the passing and impeding vehicles and grade, in that order, also have some effect. Grade is, however, not a major factor if the passing speed is less than 50 mph.

The results also indicate that the MUTCD specified requirements for marking no-passing zones are not adequate to ensure safe passing maneuvers on mountainous highways. The minimum length of 400 ft for a passing zone specified by the MUTCD may not be adequate for passing vehicles to safely complete a pass, even at a 30 mph passing speed.

RECOMMENDATIONS

It is recommended that a review of the guidelines given in the MUTCD for marking passing and no-passing zones be undertaken with the objective of updating these guidelines using results of this and other recent studies.

REFERENCES

1. Manual on Uniform Traffic Control Devices for Streets and Highways. National Advisory Committee on Uniform Traffic Control Devices, FHWA, U.S. Department of Transportation, 1978.
2. W.J. Dixon and M.B. Brown, eds. BMDP: Biomedical Computer Program P-Series. University of California Press, Los Angeles, 1979.
3. E.B. Lieberman. A Model of the Passing Maneuver on Two-Lane Rural Roads. Presented at 61st Annual Meeting of the Transportation Research Board, Washington, D.C., 1982.

Publication of this paper sponsored by Committee on Traffic Control Devices.

Abridgment

Comparative Analysis of Left-Turn Phase Sequencing

RANDY B. MACHEMEHL and ANN M. MECHLER

ABSTRACT

Guidelines for left-turn phase use do not generally include recommendations for left-turn signal phase sequence patterns. In this research, the TEXAS simulation model is employed to study the effects of various left-turn sequence patterns on traffic operations in order to establish guidelines for using most typical sequence patterns. Recent literature on the effects of left-turn sequence patterns on intersection delay and accidents is reviewed. Using vehicular delay as a basis for comparison, protected only and protected/permissive left-turn phasing with pretimed control are studied. Dual leading and dual lagging left-turn phase sequences, supplemented by permissive turning and pretimed control, are also studied. Furthermore, split, dual, and composite sequences are compared for the pretimed case. The examination of basic phase sequencing schemes under actuated signal control essentially duplicates that for pretimed control. Finally, guidelines for the implementation of phase sequence patterns are presented.

When left-turn demands approach or exceed maximum unprotected flow rates at signalized intersections, traffic control schemes are usually modified to provide protected left-turn signal phases. Guidelines for left-turn phase use do not generally provide a specific rationale for choosing among the many possible left-turn signal phase sequence patterns. The study reported in this paper (1) contains a description of the effects of various left-turn sequence patterns on left-turn as well as total intersection traffic operations. Guidelines for using most typical sequence patterns are presented.

For the purpose of this discussion the following terminology has been adopted for describing left-turn phase sequences. A protected left-turn phase is that portion of the signal cycle in which left-turn maneuvers are permitted and all conflicting maneuvers are prohibited. A permissive left-turn phase is that portion of a cycle in which left turns are permitted but only through gaps in the opposing traffic stream. A protected left-turn phase that occurs before display of the opposing through green is called a leading phase, and one occurring immediately after the opposing through green is called lagging. The term "dual left turns" is used to describe protected left-turn phases that occur simultaneously on opposing approaches of the same street. "Split phasing" is used to describe schemes in which protected left-turn phases on the same street do not occur simultaneously. Such schemes may or may not use protected left-turn phases on both approaches, and, where present, the protected lefts may occur before, after, or during through movement green indications.

PREVIOUS RESEARCH

Published research findings were reviewed to provide a background for primary data collection and analysis efforts. Five significant references, which dealt with the question of how left-turn phase sequencing affects vehicular delay and accidents, were located (2-6).

Each of these studies compared measures of vehicular delay for protected-only and protected/permissive left-turn phasing. Particular phase sequence patterns such as dual and split arrangements were not specifically addressed. Field studies conducted in Maryland, California, Florida, and Kentucky found that intersection delay was reduced when permissive left-turning supplemented the protected phase.

All of the studies compared the frequency of left-turn accidents before and after permissive phasing was installed. Experiences in the four states indicated that permissive phasing does not produce statistically significant changes in accident experience or accident severity at locations with good geometrics and approach speeds less than 45 mph.

COMPARATIVE ANALYSES

Left-turn phase sequence patterns, using both pretimed and actuated signal controllers, were compared under a variety of traffic demands. Computer simulation using the TEXAS model for intersection traffic was the primary data collection tool. Simulation provided a means of systematically examining combinations of geometry and traffic demand that were of specific interest.

Protected Versus Protected/Permissive Phasing

The first experiment conducted as part of this study compared total delay for permissive/protected and protected left-turn phases with fixed cycle and phase durations. Test conditions were imposed on a four-leg intersection in which all approaches were loaded by the same traffic volumes with a left-turn percentage of 20. The application of different traffic demands to one timing plan demonstrates the set of conditions that might exist through the various peak and off-peak hours of a typical day.

Total vehicular delay was compared for the protected-only and the permissive/protected phase patterns. The nonparametric Kolmogorov-Smirnov test was used to evaluate the statistical significance of the differences in delay statistics. The two test conditions were found to be significantly different at all volume levels, with permissive/protected phasing producing less delay. The protected/permissive sequence generally produced an 80 percent reduction in total delay to left-turn traffic.

The consistency of the delay reduction is particularly significant because opposing traffic volumes ranged from 360 vehicles per hour (vph) (80 percent of 450 vph) to 600 vph (80 percent of 750 vph). Under the low-volume condition, the unprotected left-turn capacity exceeded the demand whereas under

the 750 vph demand the unprotected capacity was less than one-third the demand. Therefore, even when a relatively small fraction of the total left-turn demand can be served by a permissive phase, large savings in left-turn delay can be expected.

Phase Sequences Under Pretimed Control

In the previous section permissive/protected phasing was shown to be generally effective in reducing vehicular delay relative to protected-only phasing. The sequence in which protected phases are provided may also have an effect on vehicular delay because most protected phase sequences may be supplemented by permissive turns.

Dual Left-Turn Phasing

Dual leading and lagging left turns were compared under conditions of protected-only left turning. This experiment used the same intersection geometry, signal timing, and traffic demands as the previous experiment. Eight hours of simulated observation time were collected for each test condition. Statistical testing of the differences in total delay between leading and lagging dual phases under protected-only phasing indicates that the two schemes are not significantly different with regard to both left-turning and total approach traffic delay.

Because of this conclusion, vehicular delay and other operational statistics were compared for leading and lagging dual left-turn phasing when both were supplemented by permissive turning. The test conditions were expanded to encompass a variety of 20 different approach traffic demands. For each case, signal phase and cycle lengths were arranged to be nearly optimal for the stated demand and at least 1 hr of simulated observation time was collected.

Nonparametric statistical testing indicates that dual leading left-turn phases produce less delay to left-turning vehicles than does dual lagging if the opposing traffic demand on two inbound lanes is less than approximately 600 vph. When opposing volumes are relatively small, significant numbers of vehicles can execute left-turn maneuvers during the permissive portion of the signal cycle. As opposing traffic volumes increase, the numbers of left turns made during permissive phases decrease until the only opportunities may occur during clearance intervals. As indicated previously, dual leading and dual lagging sequences tend to produce equivalent left-turn delays when very few turning opportunities are available during permissive green intervals. Therefore, dual leading phasing apparently provides more efficient use of unprotected left-turn phase.

Dual Versus Split Phasing

Split left-turn phasing schemes were earlier identified as any of a family of phase sequencing arrangements in which protected left-turn phases on two approaches of the same street do not occur simultaneously. Split phasing is used most effectively on a street where the maximum left-turn and through movement demands occur on the same approach. Thus, both the left-turn and through movement volumes on the opposite approach would be noncritical if both approaches were serviced by a common signal phase. This situation would be particularly appropriate for split phase sequencing with no permissive turning. Ideally, if permissive turns are to be allowed, the left-turn demand on one approach should require more processing time than the through movement, and on the opposing approach more green time should be required to process the through movement.

To compare vehicular delay resulting from dual and split phasing a series of specially designed experiments was conducted. Two nearly optimal signal timing schemes were developed for a traffic demand condition on a four-by-four intersection. In one scheme dual leading left-turn phases were imposed, and in the other, split left-turn phasing was used; permissive left turns were allowed in both. As expected, the dual left phasing produced significantly less left-turn and total intersection delay. This effect can be attributed largely to the shorter phase and cycle lengths possible with dual sequencing.

To extend the comparison and examine the effects of cycle length, the same traffic demands were used again but signal cycle lengths of 60, 80, and 150 sec were used for both the dual and the split sequences. The number of phases required for the 60-sec cycle was the same as in the optimum cycle experiment, and the results were the same.

For the 150-sec cycle, the much larger red times produced larger queues and requirements for protected left-turn phases on all four approaches. Here again, dual left-turn phasing should be better than split phasing because on both streets each approach required more time to process the main street traffic than the left-turn vehicles. The experimental results confirm this conclusion.

The 80-sec cycle, on the other hand, produced requirements for protected left-turn phases on both approaches of street A, but only one approach of street B. In this case, split phasing resulted in less total approach delay on street B, and dual phasing performed more efficiently on street A. The experiments comparing dual and split phases under pretimed control indicate that split phase sequencing should be considered as a candidate sequencing scheme where (a) the critical left-turn and through movement demands occur on the same approach, and (b) on only one approach the required left-turn processing time exceeds that for the through movement.

Phase Sequences Under Actuated Control

A testing program for left-turn phase sequencing under actuated signal control was designed to parallel that for pretimed control. A number of questions regarding detector patterns and controller timing were also studied to provide results comparable with those of the previous experiment.

Detector Configuration and Phase Timing

Sensitivity analyses along with supplemental studies of detector configurations were used to develop plans for detector configuration and phase timing. These studies, in conjunction with consideration of the traffic demands to be studied, yielded initial specifications of 90-ft-long presence detectors in the left-turn bays and across both through traffic lanes. The detectors were set back 10 ft from the stop lines.

One-second initial intervals and 1-sec vehicle extension intervals were used with 2-sec minimum greens. In all experiments permissive left turns supplemented protected left turns. Although the experiments were conducted with fully skippable phases, all phases occurred consistently on both streets. At least 20 min (and up to 90 min) of simulated observation time were collected for these traffic volume cases.

Dual Left-Turn Phasing

Operational efficiency, with vehicular delay as the principal measure of effectiveness, was compared for leading and lagging dual protected left-turn phasing when both were supplemented by permissive left turns and timed by an actuated controller.

Nonparametric statistical tests of the experimental results indicate that dual lagging left-turn phasing creates shorter cycle lengths that produce smaller delays to the dominant through movements. The main street green is used much more effectively with lagging left-turn phases because vehicles in the left-turn queue can be proceeded during gaps in the main street traffic. A leading left-turn phase, on the other hand, may process the entire left-turn queue before the main street green. Therefore, the main street green is used to process only those left-turn vehicles that arrive while it is in progress. As a result, the cycle length for the intersection is increased. In situations where the maximum phase extension is reached during the protected left-turn phase, with dual lag phasing the cycle length will be equal to or shorter than it will be with dual lead phasing.

As the statistical tests also verify, the reduction in cycle duration due to lagging phases causes a significant reduction in delay to through vehicles. Left-turn vehicles benefit from this delay reduction, but at the same time experience a delay increase caused by slower queue dissipation. Thus, left-turn vehicles may or may not benefit from either phase arrangement, depending on the left-turning and the opposing traffic demand.

For some experimental traffic arrangements lagging phases produce significantly less delay to all traffic on an approach. But when all experimental traffic demand cases were tested together, the difference was not significant. Approach delay under actuated control is dependent on the interactive performance of all vehicles using an approach and the relative efficiency and relative magnitude of each vehicular maneuver.

Dual leading left-turn phasing was compared with dual lagging for the same 20 traffic demand combinations that were examined in the corresponding pretimed experiment with four additional special cases. Because dual lag phasing generally produced shorter cycle lengths and less delay than dual lead phasing, a supplementary experiment was designed in an attempt to produce shorter cycle durations with dual lead phasing. The left-turn lane loop detectors were incrementally shortened in three tests, and a shorter vehicle extension interval was used for left-turn traffic.

Although forcing the left-turn traffic to use the permissive portion of the green signal by causing early gap-out of the protected left-turn phase causes the cycle duration to be reduced, it was never as short as with dual lag phasing. Vehicular delay was consistently less for dual lagging sequencing schemes. The dual lagging sequence was, therefore, judged to be more efficient than dual leading.

Dual Versus Split Phasing

As noted earlier, split phase timing patterns were developed for 20 traffic demand situations. Vehicular delay for through and left-turn movements was compared with the corresponding statistics gathered under dual left-turn sequencing. Results of the comparisons were virtually identical to those produced under pretimed control. Therefore, the conditions determining whether split or dual phasing should be used do not change when actuated instead of pretimed control is used.

Leading Versus Lagging Split Phasing

In cases where split left-turn sequences are selected under actuated control, the question of which left-turn movement should lead a through movement green may arise. To determine whether the leading left-turn movement performs differently than the lagging movement in a split left-turn phase arrangement, 20 traffic approach demand combinations were compared for each of the two situations.

The results indicate that there is no significant difference in delay to left-turning or to through vehicles when a lagging phase is used instead of a leading phase, even though the required phase lengths are very different. This is because the left-turn queue discharges more efficiently with a leading phase minimizing delay to individual vehicles, but it requires a longer phase to do so, causing a longer cycle duration and more delay at the intersection. On the other hand, because the lagging phase is shorter, the main street green signal must be longer to process the through vehicles that would be processed with the left-turn vehicles with a leading phase. Thus, there is no significant difference between leading and lagging phases with split left turns and actuated control.

FINDINGS AND RECOMMENDATIONS

The preceding discussion has included a comparative examination of left-turn phase sequence patterns. Random variability of generated traffic data has been considered an important aspect of the study and has been treated through multiple replication of experimental units. Comparative analyses have been developed around traffic operational data with vehicular delay as the primary measure of effectiveness. Safety-related issues have been included through a review of published safety data (2-6). Based on these analyses the following findings have been developed.

1. From a traffic operations perspective, provision of permissive left turns during the through green will always be beneficial regardless of the type of signal control or left-turn sequence pattern. Published data (2-6) indicate that safety problems associated with permissive lefts are frequently not severe. Intersection approach speeds greater than 45 mph are frequently cited as a reason for prohibiting permissive left turns.

2. There is no operational difference between dual leading and dual lagging sequences when permissive left turns are prohibited. When permissive turning is allowed, dual leading sequences produce less vehicular delay than dual lagging sequences if pretimed signal control is used. Under actuated control, dual lagging sequence patterns tend to produce less vehicular delay.

3. The choice of dual versus split phase sequence patterns is not generally affected by the type of signal controller. Split phasing will be the more efficient sequence pattern where the critical left-turn and through movement traffic demands occur on the same approach and left-turn processing time for one approach is greater than the through movement processing time.

REFERENCES

1. R.B. Machemehl and A.M. Mechler. Comparative Analysis of Left-Turn Phase Sequencing. Research Report 258-2. Center for Transportation Research, University of Texas, Austin, Nov. 1983.
2. Florida Section, ITE. Left-Turn Phase Design in Florida. ITE Journal, Institute of Transportation Engineers, Washington, D.C., Sept. 1982, pp. 28-35.
3. R.H. Mohle and T.K. Rorabaugh. A Study of Clearance Intervals, Flashing Operations, and Left-Turn Phasing at Traffic Signals, Vol. 4. Report FHWA-RD-78-49. FHWA, U.S. Department of Transportation, May 1980.
4. C.S. Wu, R.B. Machemehl, and C.E. Lee. Detector Configuration and Location at Signalized Intersections. Report 259-1F. Center for Transportation Research, University of Texas, Austin, June 1981.
5. C.J. Messer, D.B. Fambro, and D.A. Andersen. A Study of the Effects of Design and Operational Performance of Signal Systems--Final Report. Research Report 203-2F. Texas Transportation Institute, Texas A&M University, College Station, Aug. 1975.
6. K.R. Agent. An Evaluation of Permissive Left-Turn Phasing. Research Report 519. Division of Research, Kentucky Department of Transportation, Lexington, April 1979.

The contents of this paper reflect the views of the authors who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the FHWA and the Texas State Department of Highways and Public Transportation. This paper does not constitute a standard, specification, or regulation.

Publication of this paper sponsored by Committee on Traffic Control Devices.