1114

TRANSPORTATION RESEARCH RECORD

Traffic Control Devices and Rail-Highway Grade Crossings

TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL
WASHINGTON, D.C. 1987

Transportation Research Record 1114
Price: $21.50
Editor: Edythe Traylor Crump
Typesetter: Harlow A. Bickford

modes
1 highway transportation
3 rail transportation

subject areas
51 transportation safety
52 human factors
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 rail-highway grade crossings.

(Transportation research record, ISSN 0361-1981 ; 1114)
1. Traffic signs and signals. 2. Railroads—crossings. I. National Research Council (U.S.). Transportation Research Board. II. Series.
TE7.H5 no. 1114 [TE228] 380.5 s 87-28292
ISBN 0-309-04466-9 [388.3'122]

Sponsorship of Transportation Research Record 1114

GROUP 3—OPERATION, SAFETY, AND MAINTENANCE OF TRANSPORTATION FACILITIES
James I. Taylor, University of Notre Dame, chairman

Facilities and Operations Section

Committee on Traffic Control Devices
Paul H. Fowler, Automobile Club of Southern California, chairman
Earl C. Williams, Jr., Tennessee Department of Transportation, secretary
Ronald M. Cameron, Robert L. Carstens, Charles E. Dare, P. Norman Deitch, Robert E. Dewar, C. Arthur Geurts, Robert L. Gordon, Robert David Henry, Ken F. Kobetsky, Richard P. Kramer, Richard W. Lyles, C. John MacGowan, Hugh W. McGee, Zoltan A. Nemeth, Peter S. Parsonson, Ronald E. Stemmler, Dwight L. Stevens, Richard H. Sullivan, James A. Thompson, Jonathan Upchurch, Clinton A. Venable, W. Scott Wainwright, Davey L. Warren

Committee on Railroad-Highway Grade Crossings
Charles Raymond Lewis II, West Virginia Department of Highways, chairman
Otto F. Sonefeld, Atchison, Topeka, and Santa Fe Railway Co., secretary
John E. Baerwald, William D. Berg, Brian L. Bowman, Denis J. Dwyer, Lee Forrester, Bruce F. George, William H. Gossard, William J. Hedley, Robert J. Herstein, Frederick J. Kull, Richard A. Mather, William J. Matthews, G. Rex Nichelson, Jr., Paul C. Oakley, Ernie Oliphant, James M. Overton, Hoy A. Richards, Eugene R. Russell, Sr., Cliff Shoemaker, Robert E. Taggart, Jr., Earl C. Williams, Jr.

David K. Witheford, Transportation Research Board staff

Sponsorship is indicated by a footnote at the end of each paper. The organizational units, officers, and members are as of December 31, 1986.

NOTICE: The Transportation Research Board does not endorse products or manufacturers. Trade and manufacturers' names appear in this Record because they are considered essential to its object.

Transportation Research Record 1114

The **Transportation Research Record** series consists of collections of papers on a given subject. Most of the papers in a Transportation Research Record were originally prepared for presentation at a TRB Annual Meeting. All papers (both Annual Meeting papers and those submitted solely for publication) have been reviewed and accepted for publication by TRB's peer review process according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

The views expressed in these papers are those of the authors and do not necessarily reflect those of the sponsoring committee, the Transportation Research Board, the National Research Council, or the sponsors of TRB activities.

Transportation Research Records are issued irregularly; approximately 50 are released each year. Each is classified according to the modes and subject areas dealt with in the individual papers it contains. TRB publications are available on direct order from TRB, or they may 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.

Contents

Foreword

Issues of safety and the effectiveness of traffic control devices are the primary topics of this Record, which contains papers sponsored by the Committees on Traffic Control Devices and Railroad-Highway Grade Crossings.

The first six papers deal with a variety of devices. In a study of different types of delineation on horizontal curves, Zador, Stein, Wright, and Hall conclude that delineation is effective although no one type is clearly superior. Shapiro, Upchurch, Lowen, and Siaurusaitis report on an evaluation of standards in the *Manual on Uniform Traffic Control Devices*, concluding that research is a high priority need in eight different cases. Hall's study of wide edgelines on 176 miles of New Mexico highways revealed that they do not have a significant effect on the incidence of run-off-the-road accidents. Hanscom found that weight-specific signs were effective in controlling truck speeds at two of three downgrade locations, and further, that states' use of the Grade Severity Rating System improved their liability position. Seven designs of advance warning signs for median crossovers were tested by Worsey, Dare, and Schwab. The results showed that a symbolic sign was the simplest and most effective. Ullman and Dudek tested a hypothesis that speed limits lower than the 85th percentile speed might be beneficial in urban fringes; however, no significant changes in either speed characteristics or accident rates occurred.

The next six papers concern different aspects of traffic signal control. The effect of two detector patterns on three- and four-phase activated signal control at diamond interchanges was examined by Messer and Chang, who concluded that single-point detection, while better with three-phase control, was not as effective as multipoint detection in four-phase signal operation. Stone and Upchurch conducted before-and-after studies of volume and delay at signalized intersections where left-turn phasing was changed from permissive to exclusive/permissive. Improvements in left-turn movements were offset by large overall increases in delay because of longer cycles and inefficient use of green time. Bonneson and McCoy studied the procedure in the *1985 Highway Capacity Manual* for operational analysis of protected/permitted phasing at intersections, concluding that changes are needed in the analysis of left-turn lane groups. The outcome of a study on signal change intervals by Lin, Cook, and Vijayahumar was a suggestion that equations for calculating intervals should be simplified and designed to better reflect driver behavior. On the same subject, Mahalel and Prasker report that longer intervals, by creating a longer indecision zone, imply an increase in the risk of rear-end collisions. In the last paper on traffic signals, by Bullen, Hummon, Bryer, and Nekmat, a computer model (EVIPAS) is presented that is designed to optimize controller timing for activated signals at isolated intersections.

Rail-highway grade crossing controls and safety are the subjects of the last group of five papers. In a study sponsored by FHWA, Bowman found that constant warning time systems are effective, though the lowered accident rates were not significantly different statistically at the 95 percent confidence level. Halkias and Blanchard compared fixed distance and constant warning time systems at gate-protected crossings using 1975–1984 data in national files, finding that a lack of credibility in warning systems contributed to accidents. A procedure to account for causal factors in grade-crossing characteristics that will more accurately estimate the safety effects of different warning devices is offered by Hauer and Persaud. Active advance warning devices, in three configurations, were tested by Bowman in four locations with sight distance restrictions. The array with a standard 48-in. railroad advance warning sign would effectively provide the necessary warning. Last, the use of different models to establish the expected accident rate at rail-highway crossings was explored by Faghri and Demetsky, who found that the U.S. Department of Transportation (DOT) formula out-performed four other methods.

Effects of Chevrons, Post-Mounted Delineators, and Raised Pavement Markers on Driver Behavior at Roadway Curves

PAUL ZADOR, HOWARD S. STEIN, PAUL WRIGHT, AND JEROME HALL

Previous research has shown that in single-vehicle crashes drivers tend to run off the road in the direction opposite the curve; that is, they miss the curve. Examined in this study are the short- and long-term effects of commonly used curve delineation treatments on the speed and placement of vehicles traveling on curves on rural two-lane highways in Georgia (46 sites) and New Mexico (5 sites). Vehicle speed and placement distributions at sites modified with the addition of chevrons, post-mounted delineators, and raised pavement markers and unmodified control sites were compared in terms of 10th percentile, 90th percentile, mean values, and standard deviations before and after modification. The modifications tended to shift the nighttime speed distributions upward, with an average speed increase of 1 to 3 ft/sec; however, in Georgia, chevrons had little effect on speed. Overall, when chevron signs were used at night, vehicles moved away from the centerline; they moved farther away when raised pavement markers were used. In contrast, when post-mounted delineators were used, vehicles moved toward the centerline. Vehicle speed and placement variability were also slightly reduced with the use of chevrons and raised pavement markers. There was little change in the typical driver curve-following behaviors of corner cutting on curve lengthening. Few of the changes varied systemically by curve alignment or grade, and there was little evidence that short-term changes eroded over time. Although drivers did change their behavior in response to the delineation modifications, there was no clear evidence that any one of the devices is superior to the others. The primary benefit of clearly delineating curves may simply be that it helps drivers better recognize that they are approaching a curve.

Research has shown that roadway curves are often a factor in vehicle crashes, especially on rural roads *(1–3)*. During 1983, more than 25 percent of fatal highway crashes occurred on curves, and 40 percent of these crashes were also on grades *(4)*. Detailed analyses of single-vehicle crash sites show that vehicles most commonly leave the roadway on the outside of the curve; that is particularly true for left curves *(1–3, 5)*. Curves on roadways have also been shown to be more hazardous for drivers who are unfamiliar with the route *(5)*.

The most common technique used in attempting to reduce crashes on curves is to improve delineation of the roadway with roadway markings of signs. A survey of state highway agencies revealed that chevron signs, raised pavement markers, post-mounted delineators, and curve warning signs are the countermeasures most often used and judged most effective in reducing crashes (although there has been little documentation of their actual effect) *(6)*. Improving roadway delineation is also strongly supported by the U.S. Department of Transportation, which has allocated several hundred million dollars for these activities over the last decade *(7)*.

The choice of specific countermeasures at a given site should to the extent possible be guided by scientific evidence of their expected effects on crashes as well as by engineering considerations of implementation and cost. These effects could vary with road geometry and design. Because crash studies for comparing delineation modifications while controlling for other factors are time consuming and expensive, the effects of delineation modifications are more often studied in relation to the change in driver behavior they produce.

PREVIOUS RESEARCH

Before reviewing studies of driver performance with supplemental delineation systems, it is important to understand how drivers typically negotiate curves. Most drivers do not steer a circular path following the curve's radius. They tend to steer a straighter path flattening the curve until, at some point, they must steer a path that turns more sharply than the actual roadway curve *(8–10)*. This driving behavior is termed curve "lengthening." Because the actual path that drivers follow in negotiating curves is not the center of the lane, they may exceed the speed and side friction limitations for which the roadway curve was designed. If this occurs at the same point along the curve where the curvature of the driver's path is sharper than the roadway curvature, the vehicle will begin to slide laterally on the road. The question of whether curve delineation should accommodate curve lengthening or influence drivers to follow a more circular path around curves has not been satisfactorily answered. However, most researchers have interpreted a decrease in the variability in vehicle speed and lateral position to be a major benefit of improved curve delineation *(11–13)*.

Research on the effects of delineation modification on roadway curves has concentrated on studies of factors in driver perception and visibility and driver behavior. Studies of driver visibility requirements and perception of curved roadway sections have typically involved either driver simulations or driver evaluations of static pictures of curves. These studies revealed that as the range of driver visibility decreases, delineation

P. Zador and H. S. Stein, Insurance Institute for Highway Safety, Watergate 600, Suite 300, Washington, D.C. 20037. P. Wright, School of Engineering, Georgia Institute of Technology, Atlanta, Ga. 30332. J. Hall, Bureau of Engineering Research, University of New Mexico, Albuquerque, N.Mex. 87131.

becomes more important *(14)*. Also, several studies revealed that drivers have more trouble perceiving information about left curves compared with right curves *(15–18)*.

Other studies have examined the effect of both novel and conventional roadway delineation treatments on actual driving performance. Some studies of novel treatments have shown that painted markings that create an optical illusion of either increasing speed or roadway narrowing can affect driver performance and reduce crashes *(19–21)*. However, painted markings can wear rapidly and their visibility is diminished during rain. Consequently, use of these novel markings is limited.

The Federal Highway Administration (FHWA) conducted a large field evaluation of conventional and modified delineation systems, including painted centerlines and edgelines and supplemental systems such as raised pavement markers and post-mounted delineators (12). The first phase of this study evaluated driver performance at 10 curves without supplemental delineation. Vehicle placement, relative to the centerline, was measured at three points: the straightaway before the curve, the beginning of the curve, and the midpoint of the curve. Vehicle placement did not vary at the beginning of the curve compared with the straightaway, but it was significantly different at the midpoint compared with the beginning of the curve. On left-turning curves, vehicles were closer to the centerline at the curve midpoint; on right-turning curves, they were closer to the edgeline at the midpoint. Both of these behaviors are driver curve-flattening strategies.

In the second phase of the study, the speed and placement of vehicles were measured at several points along four curve sections and several tangent sections. Each section had several variations of delineation treatments. Traffic was observed at each section several days after the modifications. Nighttime midcurve speeds of vehicles traveling in both directions were lower with supplemental delineation using raised pavement markers and post-mounted delineators separately and in combination. The speeds were significantly lower (2.1 to 3.7 ft/sec) for left-turning vehicles for all the delineation modifications. Nighttime vehicle placement changes were almost always toward the edgeline for vehicles traveling in either direction. The changes were significant for raised pavement markers (and raised pavement markers in combination with post-mounted delineators), and they were larger for left-turning vehicles (0.3 to 1.1 ft). The standard deviation of vehicle placement was significantly less for three of the four supplemental delineation modifications for left-turning vehicles (0.29 to 0.16 ft/sec). The study recommended the use of raised pavement markers over post-mounted delineators on high-hazard curves because the raised pavement markers serve as both far and near delineation. It also encouraged the use of one-way raised pavement marker systems and multicolor directional coding of raised pavement markers.

Two other studies of driver performance evaluated the effects of chevron signs, different types of post-mounted delineators, and raised pavement markers; both concluded that driver performance on sharp curves was the most favorable when chevrons were used. In the first study, 36 drivers traveled a closed test track at night that had varying delineation modifications (edgelines, raised pavement markers, post-mounted delineators, and chevron signs) *(8)*. (The study was performed in Australia and drivers were on the left side of the road.) This study revealed that with chevrons drivers followed a better path around the curve (defined in terms of the ratio of the vehicle's instantaneous radius to the actual curve radius). It also revealed that drivers used a corner-cutting strategy and that chevron signs and post-mounted delineators, to some degree, facilitated this strategy. On right curves with chevrons, drivers had an average midcurve placement closest to the centerline. On left curves with chevrons, vehicle placement was not significantly different. However, with post-mounted delineators (both sides of roadway) drivers were closest to the centerline, which is contrary to the corner-cutting strategy. Higher mean vehicle speeds were found with chevrons (with and without edgelines) than with other delineations; for example, the mean speed with chevrons and no edgelines was 66 ft/sec compared with 58 ft/sec with post-mounted delineators. However, with chevrons mean nighttime speeds were not faster than daytime speeds.

The second study evaluated several types of curve delineation (chevrons, road edge delineators, and special, large striped road edge delineators) placed at five left curves in Virginia *(22)*. Speed and vehicle placement were measured at the beginning and middle of the curve. The data showed drivers were using a corner-cutting strategy, with an average 0.63-ft-difference between vehicle placement at the beginning of undelineated curves compared with the middle of these curves. The data also showed an increase in possible centerline encroachments with all of the delineation types. Although the study recommended the use of chevron signs for sharp curves, closer examination of the data indicates that it is very difficult to identify consistent differences in nighttime driver speed and placement responses to the three types of delineation.

The most important factor in evaluating delineation modifications, regardless of changes in driver behavior, is their effect on crashes. Many studies have revealed reductions in crashes and lower crash rates for roadways and curves with supplemental delineation systems *(23–27)*. However, these studies do not provide conclusive evidence of the claimed benefits because most were cross-sectional analyses and did not properly control for other factors that influence crashes such as differences in roadway design (curvature and grade) and traffic volume.

The objective of the present study was to compare changes in curve-following behavior by drivers caused by the three most common types of curve delineation devices: chevron signs, post-mounted delineators, and raised pavement markers. These devices were independently installed at curves that varied systematically in direction and degree of curvature and in steepness of grade. Previous studies have not evaluated whether the effectiveness of these devices differs by curve geometry and direction. A traffic data recorder collected vehicle speed and position data at two points along each curve section both before and after the installation of these devices. Changes in driver behavior were compared for the sites modified with the three types of delineation devices and a matched set of unmodified sites observed during similar time periods.

METHODS

Rural roadway sites were modified by Georgia and New Mexico Department of Transportation personnel following procedures in the *Manual on Uniform Traffic Control Devices* regarding the type, size, location, and spacing of the supplemental delineation *(28)*. Specific procedures used for modifying the sites with the three types of delineation are described next. All sites, including the comparison sites, had edge-line markings.

Raised pavement markers. Standard 4 x 4-in. amber Stimsonite markers were installed at the selected sites on both sides of the double yellow centerlines. Reflectorized Type 1 markers, visible to both directions of traffic, were installed with two-part epoxy in a sloped 52-in.-groove (26 in. long in each direction) so that the top of the marker was flush with the original surface. The markers were usually spaced 80 ft apart; along the sharper curves, where at least three markers could not be seen at one time, they were spaced 40 ft apart. The markers were installed throughout the length of the curve. The typical cost to modify a site was approximately $250.

Post-mounted delineators. Standard 3-in.-diameter round, white Stimsonite delineators were installed on metal posts along the outside of the curves. The delineators were installed on both sides of the posts in order to be visible to drivers traveling in both directions. The delineators were placed approximately 4 ft above the near roadway edge and 7 ft away from the edge of the pavement. Where shoulders were less than 7 ft wide, the delineators were placed as close as practicable to the shoulder edge. The delineators were spaced so that drivers would see at least three delineators simultaneously. The typical cost to modify a site was approximately $200.

Chevron signs. Standard 18 x 24-in. chevron alignment signs were placed along the outside of the curves in order to be visible to drivers traveling in both directions. The signs were positioned so that motorists would always have at least three in view. The signs were offset 7 ft away from the pavement or as close as practicable to the shoulder edge when the shoulder was less than 7 ft wide and were mounted at a height of approximately 7 ft. The typical cost to modify a site was $300 to $400.

Traffic Data Recorder

A special traffic data recorder (TDR) was constructed by the University of New Mexico Engineering Research Institute to measure the speed and placement of vehicles as they traveled along the road. The TDR consisted of an arrangement of electronic tapeswitches on the roadway and a Rockwell AIM-65 microprocessor with a printer that interpreted and printed the actuations of the tapeswitches. The first tapeswitch (spanning the road in the direction of travel) alerted the TDR of an approaching vehicle and counted the total traffic in both directions. The next two tapeswitches were placed a fixed distance apart to serve as a "trap" for measuring vehicle speed. Once the speed was known, the placement of the vehicle's right front tire could be computed from a fourth tapeswitch placed at a 45 degree angle to the second and third tapeswitches. Vehicle position, speed, and placement, and the time of the vehicle and traffic counts were printed onto a paper tape after the vehicle cleared all the tapeswitches.

Preliminary testing of the TDR by placing it at several points along a curve indicated that, at about 100 ft before the beginning of the curve, drivers had yet to begin adjusting for the upcoming curve. Over the next 200 ft most of the change in placement occurs and the vehicle path is defined. Several studies have examined the speed and placement of vehicles at the center of the curve; however, because drivers tend to flatten out curves, the major effects of the different delineation treatments might be to influence the initial adjustments drivers make when they begin to negotiate the curve. Therefore, two TDRs were set up for each day/night observation period; one 100 ft before the beginning of the curve and one 100 ft after the beginning of the curve.

Only vehicles that were isolated from all other traffic, either following or oncoming, for at least 2.5 sec were analyzed in t this study. At 70 ft/sec, the approximate average speed encountered in the study, the measurement error for an individual vehicle had a standard deviation of about 1.3 percent or 0.9 ft/sec. Thus, the standard deviation for an average speed based on 100 individual measurements was about 0.1 ft/sec. The comparable figure for the standard deviation of an average placement was about 0.01 ft.

Experimental Design

There were 46 observation sites in Georgia and 5 in New Mexico. All sites were located on two-lane rural highways. The sites in Georgia represented a nearly complete factorial design with four factors: modification (M), direction of turn (T), vertical alignment (G), and sharpness of curve (C). (Two sites had to be eliminated from the analyses because of modifications that were not part of the experiment.) There were four levels of treatment (control, chevron, post-mounted delineator, and raised pavement marker), two directions of turn (left and right), three types of vertical alignment (grade < -2 percent or down, -2 percent $\leq$ grade ≤ 2 percent or level, and 2 percent $<$ grade or up), and two levels of sharpness of curve (less sharp or more sharp within the grade and turn class). Because only a small number of sites in New Mexico were available for experimentation, only chevrons were tested.

The roadway characteristics of these sites by direction of curve and modification type are given in Table 1. The data indicate that there were some differences in the physical layout of the roadways. For example, the average of the superelevation rates at the unmodified left curves was about one-half of the average rates at the modified sites. The average speed limits

TABLE 1 AVERAGE ROADWAY CHARACTERISTICS OF GEORGIA SITES BY MODIFICATION TYPE AND DIRECTION OF CURVE

Modification	Lane Width (ft)	Shoulder Width (ft)	Superelevation at Curve (%)	Speed Limit (mph)
(a) *Left Curves*				
No modification	11.8	9.1	2.5	52
Chevron signs	12.2	11.4	5.2	51
Post-mounted delineators	11.9	9.0	6.4	45
Raised pavement markers	11.8	7.5	4.8	48
(b) *Right Curves*				
No modification	12.2	14.8	4.9	53
Chevron signs	12.0	10.1	5.3	51
Post-mounted delineators	12.3	6.4	6.9	53
Raised pavement markers	12.2	11.7	5.3	52

varied by about 14 percent for left curves and 4 percent for right curves. In addition, the sites with chevrons and raised pavement markers had the fewest curve warning signs on the approach to the curve, whereas only one unmodified site was without any type of signing (e.g., curve warning, speed limit). These relatively minor differences are unlikely to influence the before-and-after comparisons of the modification effects.

Observations were taken at each modified and control site shortly before and shortly after (several weeks) the modifications were put in place. To determine the long-term effects of the modifications, a third set of observations were taken approximately 6 months after the modifications at about one-third of the Georgia sites and at all New Mexico sites. During each of the three observation periods, data were recorded for about 100 to 150 vehicles during the day and for a similar number of vehicles at night (defined as the time of sunset).

Statistical Analysis

The effects of the modifications on curve-following behavior were investigated using seven variables:

- $V1$ = approach speed, measured 100 ft upstream from the beginning of the curve, ft/sec;
- $V2$ = curve speed, measured 100 ft downstream from the beginning of the curve, ft/sec;
- $D1$ = vehicle placement 100 ft upstream from curve, distance from centerline of road to right wheel of vehicle measured in conjunction with $V1$, ft;
- $D2$ = same as $D1$ but measured in conjunction with $V2$, ft;
- DE = estimated deceleration, computed as $(V2^2 - V1^2)/400$, ft/sec^2;
- D = average placement, computed as $(D2 + D1)/2$, ft; and
- D = change in placement between the two traps, computed as $(D2 - D1)$.

The distribution of each of the variables was summarized using four statistics: mean, standard deviation, 10th percentile, and 90th percentile. These statistics were estimated for day and night data separately by site and period of observation.

Changes in these statistics before modification compared with the first period after modification were analyzed using the general linear model (GLM) procedure developed by the SAS Institute *(29)*. The same model was used to analyze changes in all of the variables. In this model, the dependent variable, for example, the average approach speed ($MV1$), was represented in terms of main effects for modification (M), turn direction (T), vertical alignment (G), sharpness of curve (C), and the interactions of T, G, and C with the modification factor (M):

$$MV1_{mtgc} = A + B_m + C_t + D_g + E_c + F_{mt} + G_{mg} + H_{mc} + Error_{mtgc}$$

where

m = 0 for no modification,
= 1 for chevrons,
= 2 for post-mounted delineators,
= 3 for raised pavement markers;
t = 1 for left curves,
= 2 for right curves;
g = 1 for downhill grades,
= 2 for level grades,
= 3 for uphill grades; and
c = 1 for less sharp curves,
= 2 for more sharp curves.

The short-term modification effects due to chevrons in New Mexico were tested for statistical significance by using a t-test for comparing the changes between corresponding before-modification and after-modification site averages. This method of paired t-tests was also used to compare short-term and longer effects by modification groups in both Georgia and New Mexico.

TABLE 2 AVERAGE NIGHTTIME SPEED AND PLACEMENT VALUES BEFORE MODIFICATION, BY SITE CHARACTERISTICS

		Vehicle Speed (ft/sec)				Vehicle Placement (ft)			
		100 ft Before Curve		100 ft Into Curve		100 ft Before Curve		100 ft Into Curve	
		$V1$	SD	$V2$	SD	$D1$	SD	$D2$	SD
(a) By Modification Type									
No modification	(N=12)	76.5	4.0	74.1	4.3	7.4	1.1	8.0	1.0
Chevrons, Ga.	(N=10)	70.9	5.2	69.1	5.6	7.8	0.8	7.8	1.3
Chevrons, N. Mex.	(N=5)	73.6	5.3	71.9	6.0	7.9	1.3	7.6	2.4
Post-mounted delineator	(N=12)	74.0	4.8	71.5	5.2	7.4	1.0	7.8	1.0
Raised pavement markers	(N=12)	72.2	7.7	69.4	9.2	7.5	0.9	7.6	1.0
(b) By Grade									
Uphill	(N=15)	73.0	5.6	69.9	6.8	7.5	0.7	8.0	1.2
Level	(N=19)	73.7	5.9	71.6	5.9	7.6	0.9	8.0	1.3
Downhill	(N=17)	73.7	6.0	71.9	6.8	7.6	1.2	7.4	1.1
(c) By Curvature (Georgia data only)									
Left—Moderate	(N=12)	75.5	5.4	73.1	5.8	8.1	0.6	7.4	0.7
Left—Sharp	(N=11)	72.8	6.1	70.6	7.2	8.1	0.6	6.7	1.2
Right—Moderate	(N=11)	76.0	3.8	73.8	4.2	7.0	0.7	8.4	0.7
Right—Sharp	(N=12)	70.1	6.4	67.2	7.3	6.7	0.8	8.2	0.9

NOTE: SD = standard deviation.

RESULTS

Initial Vehicle Speed and Placement

Speed and placement observations for $V1$, $V2$, $D1$, and $D2$ are summarized before modification for the night data given in Table 2. All values in this table are based on the average values of the variables for the sites. Both the average speed and the average vehicle placement varied relatively little among the different modification groups (Table 2a). The range for approach speed was from 70.9 ft/sec to 76.5 ft/sec and for curve speed from 69.1 ft/sec to 74.1 ft/sec. For all modification groups, the curve speeds were a few feet per second below the approach speeds. Average vehicle placements ranged from 7.4 ft to 7.9 ft at the first speed trap and from 7.6 ft to 8.0 ft at the second trap.

The average speed and placement at the first trap varied little by grade (Table 2b). However, at the second trap, 100 ft into the curve, the average speeds were more than 1 ft/sec lower at uphill curves than at level or downhill grades, and vehicles moved away from the centerline by about 0.5 ft between the two traps at uphill and at level curves but drew closer to the centerline by 0.2 ft at downhill curves.

For both left and right curves and at both speed traps, sharper curves had lower average speeds than less sharp curves. The average differences were about 3 ft/sec for left curves and about 6 ft/sec for right curves (Table 2c). The average vehicle placement relative to the centerline was reduced by about 1 ft for left curves and increased by about 1.4 ft for right curves, which indicates a considerable amount of corner cutting or curve flattening among drivers.

Short-Term Effects of Roadway Delineation Modification—Georgia Data

To demonstrate the effects of the modifications for the Georgia data, the statistically significant changes are summarized in Table 3a for standard deviations and in Table 3b for the 10th percentiles (L), means (M), and 90th percentiles (H).

Figure 1 shows the means of the speed and placement observations before and after the modifications by time of day, direction of turn, and type of modification. Figures 1a–1c show results for chevrons, post-mounted delineators, and raised pavement markers; Figure 1d shows the data for the unmodified sites. Before-and-after speed averages are shown as bar charts for approach ($V1$) and curve ($V2$) speeds. Before-and-after vehicle placement averages 100 ft ahead of the curve ($D1$) and 100 ft into the curve ($D2$) are shown on a pair of reference lines representing the roadway section 7 and 8 ft to the right of the centerline. On all graphs the solid lines represent observations before modifications and the broken lines indicate observations after the modifications. Note when referring to these figures that the scales used for vehicle placement and speed are arbitrary. The reader will find it helpful to refer to these figures throughout the subsequent description of the results.

The presence of corner-cutting or curve-flattening behavior is clearly shown for all conditions in Figure 1. For example, on the approach to right curves, drivers are much closer to the centerline than they are at the trap 100 ft into the curve. By shifting their initial position away from the centerline and angling their vehicles in the direction of the curve, drivers reduce (or cut) the sharpness of the curve that the vehicle will travel. This maneuver also lengthens the portion of the roadway on which the vehicle travels a curved path.

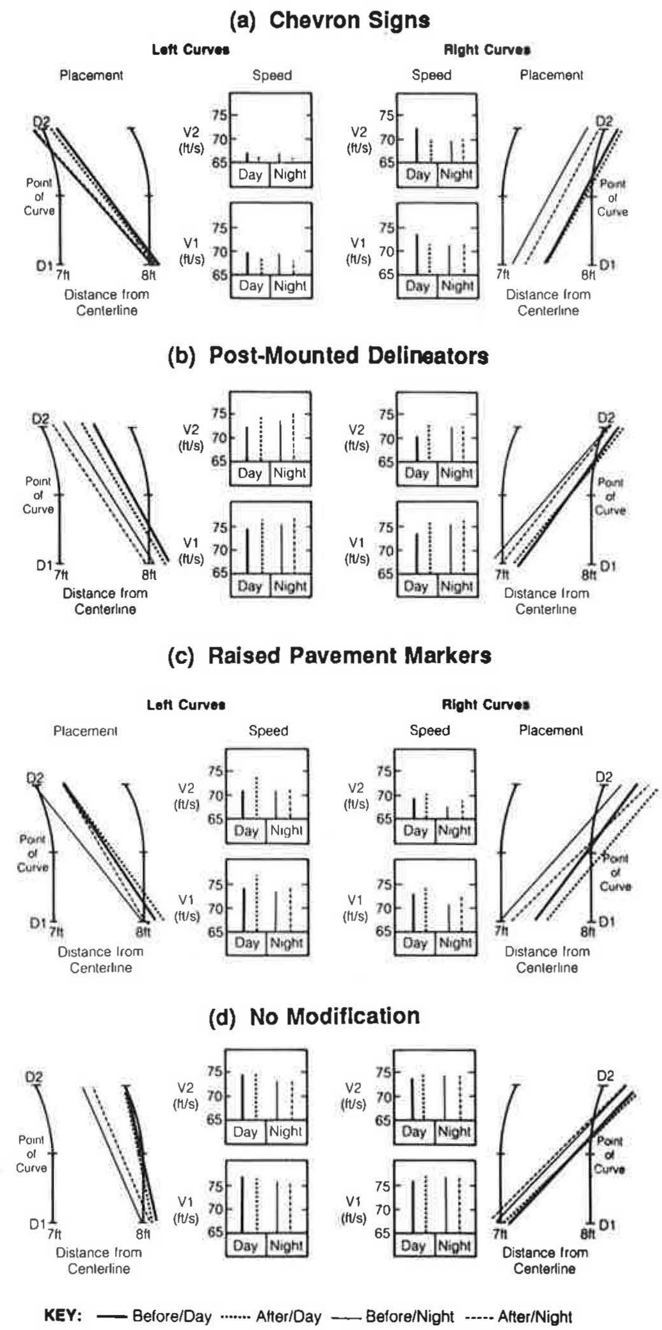


FIGURE 1 Mean speed and placement observations before and after modification by time of day, direction of turn, and treatment.

On average, after modification vehicle paths were shifted away from the centerline on right and left curves with raised pavement markers and chevrons and toward the centerline on right curves with post-mounted delineators. Placement changes were largest with raised pavement markers. Under nearly all conditions, vehicles traveled slower and nearer the centerline at night than during the day, and curve speeds ($V2$) were typically lower than approach speeds ($V1$). However, compared to the unmodified curves, nighttime speeds increased with post-mounted delineators and raised pavement markers, but they were not consistent with chevrons.

Changes in Standard Deviation

As seen from Table 3a, the standard deviations varied significantly by modification at night for the changes in two of the placement measures ($D2$ and D) and for the change in the deceleration (DE). For daytime observations, no significant main effects were found. The standard deviation in the placement 100 ft into the curve was reduced by about 0.1 ft with chevrons and raised pavement markers and increased by about 0.1 ft with post-mounted delineators. Changes in the average placement were similar in direction and in magnitude to those at the second trap. Estimated average short-term changes in the standard deviation of the deceleration showed a reduction of almost 0.2 ft/sec^2 for post-mounted delineators and increases of about 0.15 ft/sec^2 for chevrons and raised pavement markers. Also, at night all modifications, particularly chevron signs, resulted in an overall reduction in the standard deviations of curve speeds; however, this effect failed to reach the conventional level of statistical significance ($F_{3,26}$, $p = 0.076$). There was no systematic pattern of significant changes in the standard deviations associated with the modification by alignment interactions.

Changes in Mean and Percentiles

As can be seen from Table 3b, the estimated changes in mean and 90th percentile speeds exhibited significant variations by type of modification during both time periods and at both speed traps. The corresponding estimates are plotted in Figure 2 for the night observations only; the daytime changes were similar. The estimated mean approach speeds (left side of Figure 2) were reduced by about 0.6 ft/sec with chevrons, increased by about 1.1 ft/sec with raised pavement markers, and increased

TABLE 3 SHORT-TERM EFFECTS OF ROADWAY DELINEATION MODIFICATIONS ON VEHICLE SPEED AND PLACEMENT IN GEORGIA

Speed and Placement Variable	Time of Day	Modification Main Effects	Interactions: Curve Direction	Interactions: Grade	Interactions: Curve Sharpness
(a) Changes in Standard Deviation					
$D1$	Day	–	–	*	–
	Night	–	–	–	–
$D2$	Day	–	*	*	–
	Night	*	–	–	–
$V1$	Day	–	*	*	*
	Night	–	–	–	–
$V2$	Day	–	–	*	*
	Night	–	–	–	–
DE	Day	–	–	–	*
	Night	*	–	–	*
ΔD	Day	–	–	–	–
	Night	–	*	–	–
D	Day	–	–	*	–
	Night	*	–	–	–
(b) Changes in 10th Percentile (L), Mean (M) and 90th Percentile (H)					
		L M H	L M H	L M H	L M H
$D1$	Day	– – –	– – –	– – –	– – –
	Night	* * –	– – –	* * *	– * *
$D2$	Day			– –	– – –
	Night	* * –	– – –	– – –	– – –
$V1$	Day	– * *	* – –	– – –	– – –
	Night	– * *	– – –	– – –	– – –
$V2$	Day	– * *	– – –	– – –	– – –
	Night	– * *	– – –	– – *	– – –
DE	Day	– – –	– – –	– – –	– – –
	Night	– – –	– – –	– – –	– – –
ΔD	Day	– – –	– – –	– – –	– – –
	Night	– – –	– – –	– – –	– – –
D	Day	– – –	– – –	– – *	– – –
	Night	* * –	– – –	– – –	– – *

NOTES: An asterisk indicates F statistic is significant at 0.05 level; a dash indicates it is not. See section on Statistical Analysis for definition of variables. Briefly, $D1$ and $D2$ are distances from centerline 100 ft before and after the curve; $V1$ and $V2$ are the corresponding speeds; DE is deceleration, D is the average placement; and ΔD is the change in placement.

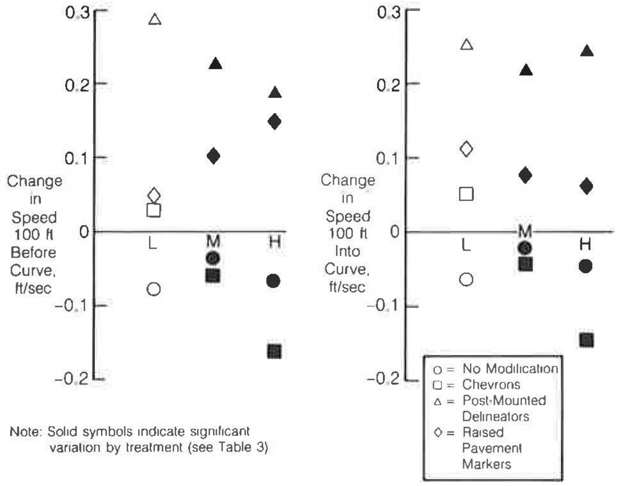


FIGURE 2 Estimated short-term changes in the 10th percentile (*L*), mean (*M*), and 90th percentile (*H*) values of speed measurements by treatment type at night for Georgia sites.

by about 2.3 ft/sec with post-mounted delineators. For the modifications at which the mean speed was reduced (particularly chevrons), the reductions were even greater for the 90th percentile speed. The pattern of changes in curve speeds (right side of Figure 2) is similar to the pattern of changes in approach speeds. This is consistent with the finding that there were no statistically significant changes in the corresponding deceleration variables.

Table 3b also shows that the estimated changes in mean and 10th percentile vehicle placement exhibited significant variations by type of modification during both time periods and at both speed traps. Figure 3 shows these estimated changes for night observations. The largest changes in vehicle placements occurred following the installation of raised pavement markers. On average, the 10th percentiles of the placement distributions shifted about 0.3 ft away from the centerline at the first trap and about 0.7 ft at the second trap. The corresponding changes in the mean placements were about 0.4 ft and 0.7 ft, respectively.

Chevrons also caused the vehicle placement distributions to shift away from the centerline, but these shifts were generally less pronounced. Overall, post-mounted delineators shifted the placement distributions toward the centerline, but the average magnitude of these shifts was quite small except for the 10th percentile placement value of about −0.3 ft at the second speed trap. At sites with post-mounted delineators vehicle placement changes varied more by direction of curve than with the other treatments, but the difference did not reach the conventional level of significance. On left curves vehicles moved toward the centerline and on right curves they moved away from the centerline. In both cases, this movement was away from the delineators, which were on the outside of the curve.

The pattern of changes in average placements were similar to those shown in Figure 3 for the 10th percentile, mean, and 90th percentile and are not displayed separately. The relative placement changes over the speed trap were not pronounced enough to cause significant changes in any of the statistics based on $D = D2 - D1$.

Modification-by-Curve-Alignment Interactions

In addition to the main modification effects, some of the modification-by-alignment interactions are statistically significant. As an illustration, Figure 4 shows the modification effect on mean placement by grade of curve (left figure) and by sharpness of curve (right figure). Neither these nor any of the other significant interactions appear to have a clear interpretation.

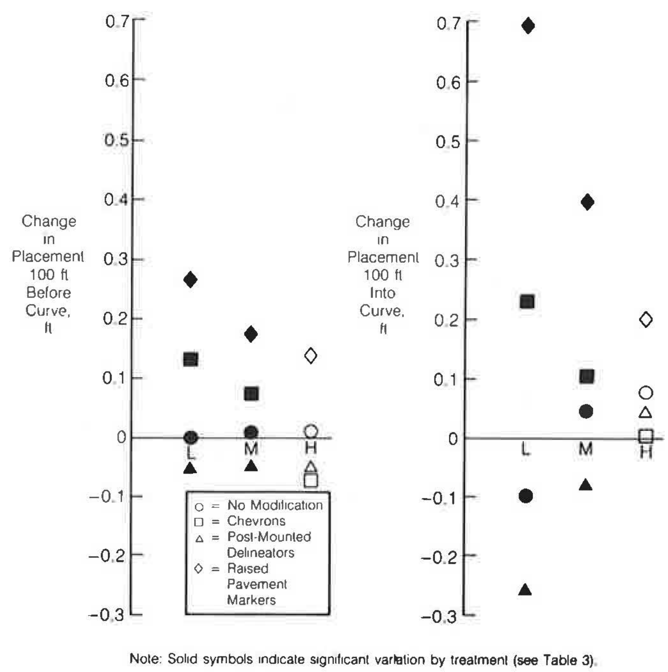


FIGURE 3 Estimated short-term changes in 10th percentile (*L*), mean (*M*), and 90th percentile (*H*) values of vehicle placement measurements by treatment type at night for Georgia sites.

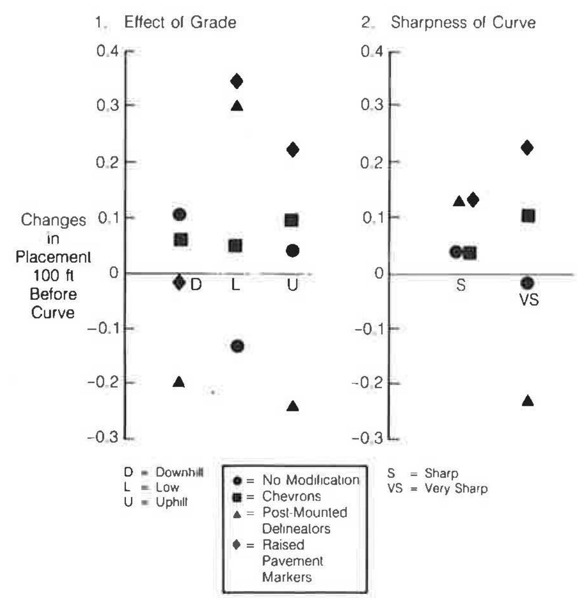


FIGURE 4 Estimated short-term changes at night in mean vehicle placement 100 ft before the curve by treatment and geometric condition at Georgia sites.

Short-Term Effects of Chevrons—New Mexico Data

The New Mexico data were limited to five sites modified with chevron signs. The short-term effects were to increase both speeds, $V1$ and $V2$, at night. There was a 3.2 ft/sec increase in approach speed that was statistically significant based on paired t-test comparisons ($t = 3.27$, $p = 0.03$), and a 2.6 ft/sec increase in curve speed that was not ($t = 2.26$, $t = 0.09$). (It should be recalled that, overall, speeds in Georgia did not increase as a result of the use of chevrons.) At night vehicles moved away from the centerline after the installation of chevron signs; however, these changes were not statistically significant.

Long-Term Effects of Delineation Modification—Georgia and New Mexico Data

Finally, to assess the long-term effects of the modifications, the averages of short- and long-term changes in the two speed measurements by type of modification are shown in Figure 5. For the Georgia data, the results are based on only those sites where three sets of measurements were taken; there were four such sites per treatment group. All five sites in New Mexico had three sets of measurements. The corresponding data for placement averages are shown in Figure 6. (Note that the results shown in Figures 5 and 6 are not directly comparable with the results based on all Georgia survey sites discussed earlier.)

Comparisons of long- and short-term differences in the speed and placement averages show three situations that were statistically different. Average curve speeds for the untreated group of curves differed by 1.7 ft/sec ($t = 3.85$, $p = 0.03$), and for raised pavement markers approach speed increased by 2.3 ft/sec ($t = 4.4$, $p = 0.02$) and curve speed increased by 2.0 ft/sec ($t = 7.2$, $p = 0.01$).

SUMMARY AND DISCUSSION

The short- and long-term effects of three commonly used delineation modifications on curve-following behavior on rural roads in Georgia and New Mexico were examined. The principal findings of this research are (a) all delineation modifications affected driver behavior at night as measured by speed and placement, (b) few systematic differences were found in the effects by type of modification or roadway alignment, and (c) these effects did not change over time. The presence of delineation modifications significantly influenced vehicle speeds and placements compared to measurements taken at unmodified sites, but there was no convincing evidence to support a preferential choice of any of these devices. There were changes at the unmodified sites although they were almost always small and unsystematic compared with those at the modified sites. The fact that most run-off-the-road crashes occur when the driver misses the curve implies that the driver has failed to control the vehicle's speed or position, or both, in sufficient time to safely negotiate the curve. The main effect of any of these delineation modifications may simply be that the driver is alerted earlier that a curve is ahead.

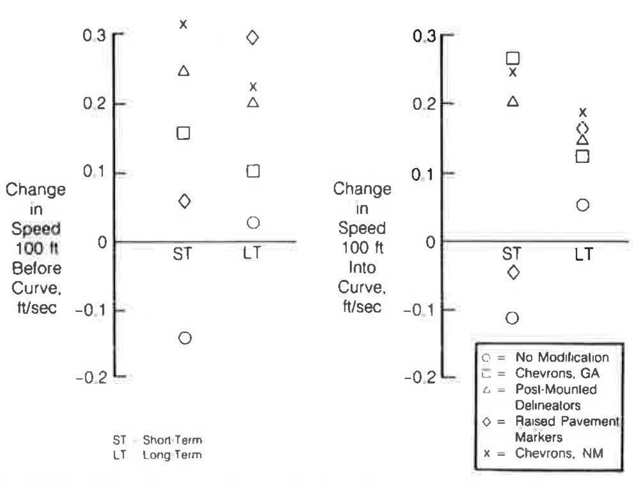


FIGURE 5 Comparison of short-term and long-term changes in vehicle speed.

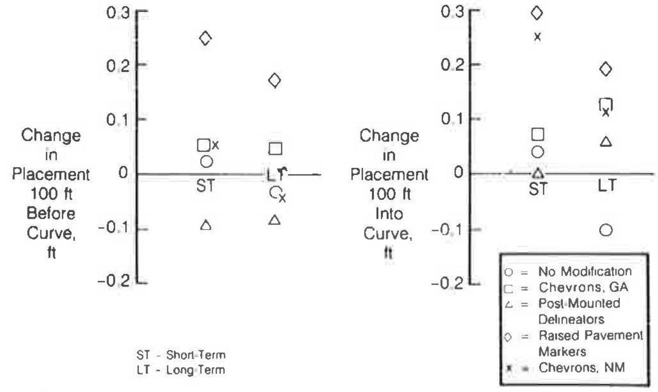


FIGURE 6 Comparison of short-term and long-term changes in vehicle placement.

The short-term results indicated that installation of post-mounted delineators produced the largest speed increases (about 2 ft/sec to 2.5 ft/sec at night). Speed increases of about 1 ft/sec at night occurred with raised pavement markers. The results for chevrons were not consistent; speed decreased by about 0.5 ft/sec at night in Georgia but increased by about 3 ft/sec in New Mexico. The long-term measurements provided no evidence of the erosion of any of these short-term speed changes.

A recent survey of state highway officials revealed that speed reductions are commonly believed to be the best surrogate for evaluating the effectiveness of measures taken to prevent run-off-the-road crashes *(6)*. On this basis, none of these devices could be advocated for use as countermeasures. However, the present study shows that, although night speeds increase with post-mounted delineators and raised pavement markers (and with chevrons in New Mexico), the resulting speeds almost always remain below the daytime speeds. It could be argued that these speed increases simply reflect driver adaptation to increased information about nighttime rural roadway conditions and are, therefore, advantageous.

Vehicle placements at night were also affected by the modifications. Generally, vehicle paths were shifted away from the centerline on curves where raised pavement markers and chevrons were installed and toward the centerline when post-mounted delineators were used, although the latter effect was present only for right curves. Changes in vehicle placement were largest at sites with raised pavement markers. The magnitudes of the shifts were about the same at both speed traps except where raised pavement markers were used when the shift at the second trap exceeded the shift at the first trap by about 0.2 ft regardless of the direction of the turn. These results can be interpreted in terms of changes in corner-cutting behavior.

For left curves, corner cutting involves first a shift away from the centerline before the curve; for right curves, the first shift is toward the centerline. The direction of these shifts is then reversed as the vehicle travels through the curve (Figure 1). Thus, the modifications had no effect on corner-cutting behavior except when raised pavement markers were used. On right curves, raised pavement markers slightly increased corner cutting during the day and at night. On left curves raised pavement markers reduced corner cutting at night and increased it during the day.

All the present and earlier studies clearly demonstrated drivers' preference for the corner-cutting strategy. Corner cutting can reduce the lateral acceleration through a curve and thereby reduce peak friction demand, but it may also bring vehicles closer to the roadway boundaries and reduce their margin of safety. However, to assess the relative importance of these factors requires the use of crash data, and previous analyses of the relation between crash frequency and implementation of delineation devices have been unable to quantify their effects or examine potential differences among devices.

The size of the changes in vehicle speeds and placements measured in this study compares well with results from other studies, but there are some inconsistencies in the directions. For example, the FHWA study *(12)* revealed that midcurve speeds were often significantly lower with raised pavement markers and post-mounted delineators, whereas in the present study speeds increased with the installation of these devices, particularly post-mounted delineators. However, both studies revealed that raised pavement markers had the largest effect on vehicle placement—vehicles moved away from the centerline. The Australian study revealed that speeds were significantly higher with chevron signs (8), but in the present study only the New Mexico sites experienced a significant short-term speed increase.

In conclusion, the results of this study provided strong evidence that supplemental delineation treatments are effective for warning drivers of approaching curves.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the cooperation of representatives of the Georgia Department of Transportation and the New Mexico State Highway Department for their overall contributions to this project. Without that cooperation, the project would not have been possible. The authors particularly appreciate the support of Archie Burnham, Traffic and Safety Engineer, and Dick Graves, Transportation Engineer, Georgia Department of Transportation, and that of Louis Medrano in New Mexico.

Special thanks are due Steven Pickett of the New Mexico Engineering Research Institute who designed and built the TDR equipment and to Chris Swenson who played a key role in the collection of the Georgia data.

Finally, appreciation is expressed to Randy Whitfield for his help with the data processing and to Sharon Rasmussen for her editorial assistance.

REFERENCES

1. P. H. Wright and L. S. Robertson. Amelioration of Roadside Obstacle Crashes. *Compendium of Technical Papers.* Institute of Transportation Engineers, Washington, D.C., 1979.
2. P. H. Wright and P. Zador. Study of Fatal Rollover Crashes in Georgia. In *Transportation Research Record 819*, TRB, National Research Council, Washington, D.C., 1982, pp. 8–17.
3. J. W. Hall and P. Zador. Survey of Single Vehicle Fatal Rollover Crash Sites in New Mexico. In *Transportation Research Record 819*, TRB, National Research Council, Washington, D.C., 1982, pp. 1–8.
4. *Fatal Accident Reporting System*, 1983. NHTSA, U.S. Department of Transportation, 1983.
5. H. H. Bissell, G. B. Pilkington, J. M. Mason, and D. C. Woods. *Roadway Cross Section and Alignment in Synthesis of Safety Research Related to Traffic Control and Roadway Elements.* Report FHWA-TS-82-232. FHWA, U.S. Department of Transportation, 1982.
6. P. H. Wright, J. W. Hall, and P. L. Zador. Low-Cost Countermeasures for Ameliorating Run-Off-the-Road Crashes. In *Transportation Research Record 926*, TRB, National Research Council, Washington, D.C., 1983.
7. *The 1985 Annual Report on Highway Safety Improvement Programs: Report to Congress.* Office of Highway Safety, U.S. Department of Transportation, 1985.
8. I. R. Johnson. The Effects of Roadway Delineation on Curve Negotiation by Both Sober and Drinking Drivers. *Australian Road Research Board*, No. 128, 1984.
9. C. R. Swenson. *A Study of Driver Behavior in Horizontal Curves* (unpublished M.S. thesis). School of Civil Engineering, Georgia Institute of Technology, Atlanta, 1985.
10. J. R. McLean. Speeds on Curves: Side Friction Factor Considerations. *Australian Road Research Board*, APR 126, March 1983.
11. N. D. Nedas and S. Luminello. *Field Study of the Influence of Post-Delineations on Driver Performance in Various Weather and Light Conditions.* Potters Industries, Hasbruck Heights, N.J., 1982.
12. W. A. Stimpson, H. W. McGee, W. K. Kittelson, and R. H. Ruddy. *Field Evaluation of Selected Delineation Treatments on Two-Lane Rural Highways.* Report FHWA 77-118. U.S. Department of Transportation, 1977.
13. J. I. Taylor, H. W. McGee, E. L. Sequin, and R. S. Hostetter. *NCHRP Report 130: Delineation Systems.* HRB, National Research Council, Washington, D.C., 1972.
14. R. W. Allen, J. F. O'Hanlon, D. T. McRuer, et al. *Driver's Visibility Requirements for Roadway Delineation, Vol. I: Effects of Contrast and Configuration on Driver Performance and Behavior.* Report FHWA-RD-77-165. U.S. Department of Transportation, 1977.
15. B. N. Fildes and T. J. Triggs. Effects of Road Curve Geometry and Approach Distance on Judgements of Curve Exit Angle. *Proc., Australian Road Research Board, Vol. II, Part 2*, Nunawading, 1982.
16. Z. A. Nemeth, T. H. Rockwell, and G. L. Smith. *Recommended Delineation Treatments at Selected Situations on Rural State Highways, Part 2.* Report FHWA-OH-85-002. U.S. Department of Transportation, 1985.
17. T. J. Triggs, J. W. Meehan, and W. G. Harris. *A Laboratory-Based Study of the Effect of Road-Side Post Frequency and Location on Curve Direction Estimation.* AIR 355-1. Australian Road Research Board, Nunawading, 1983.
18. G. R. Watts and A. R. Quimby. *Aspects of Road Layout that Affect Drivers' Perception and Risk Taking.* Transport and Road Research Laboratory Report 920. United Kingdom Department of the Environment, 1980.
19. R. O. Helliar-Symons. *Yellow Bar Experimental Carriageway Markings: Accident Study.* Transport and Road Research Labora-

tory Report 1010. United Kingdom Department of the Environment, 1981.
20. C. G. Perez. *Effect of Advance Information Markings on the Safety of Curves* (unpublished M.S. thesis). School of Civil Engineering, Georgia Institute of Technology, Atlanta, 1979.
21. D. Shinar, T. H. Rockwell, and J. A. Maleck. The Effects of Changes in Driver Perception on Rural Curve Negotiations. *Ergonomics*, Vol. 23, No. 3, 1980, pp. 263–275.
22. B. E. Jennings. *Evaluation of Rural Highway Curve Delineation Signs* (unpublished M.S. thesis). School of Engineering, University of Virginia, Charlottesville, Jan. 1984.
23. S. Bali, R. Potts, J. A. Fee, J. I. Taylor, and J. Glennon. *Cost-Effectiveness and Safety of Alternative Roadway Delineation Treatments for Rural Two-Lane Highways, Vol. II.* Report FHWA-RD-78-51. U.S. Department of Transportation, 1978.
24. C. L. Kugle, O. J. Pendleton, and M. S. Von Tress. *An Evaluation of the Accident Reduction Effectiveness of Raised Pavement Markers.* Texas Transportation Institute, College Station, 1984.
25. W. Niessner. *Post-Mounted Delineators.* Report FHWA-TS-83-208. FHWA, U.S. Department of Transportation, 1983.
26. W. C. Taylor and T. J. Foudy. Ohio's Curve Delineation Program—An Analysis. *Traffic Engineering*, Vol. 10, No. 10, 1966, pp. 41–45.
27. P. L. Zador, P. H. Wright, and R. Karpf. *Effect of Pavement Markers on Nighttime Crashes in Georgia.* Insurance Institute for Highway Safety, Washington, D.C., 1982.
28. *Manual on Uniform Traffic Control Devices.* FHWA, U.S. Department of Transportation, 1978.
29. *SAS User's Guide: Statistics, 1982 Edition*, SAS Institute, Cary, N.C., 1982.

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

Identification of Needed Traffic Control Device Research

PHILLIP S. SHAPIRO, JONATHAN E. UPCHURCH, JOHN LOEWEN, AND VIC SIAURUSAITIS

The 50-year evolutionary development of the current *Manual on Uniform Traffic Control Devices* (MUTCD) has resulted in many traffic control device standards that are based on subjective opinion. As vehicle design, the driver population, and society's demand for a safer highway system change, many of these traffic control device standards need reexamination. A research study was conducted to identify those MUTCD standards that (a) lack a research basis, or are in conflict with research findings; and (b) would likely benefit from research and scientific investigation. Nearly all MUTCD standards were evaluated. Identification of those MUTCD standards having the greatest need for additional research was achieved through (a) evaluation of selected MUTCD standards by the project team and a panel of traffic engineering practitioners, and (b) evaluation of relevant previous traffic control device research. Seventeen MUTCD standards were identified as having a significant need for additional research. Eight areas were recommended as having high priority for future traffic control device research. To provide a tool for future research and to serve as an aid to ongoing development of the MUTCD, a computerized data base management system was created. It includes documentation of previous traffic control device research as it relates to each standard within the MUTCD.

The *Manual on Uniform Traffic Control Devices (MUTCD) for Streets and Highways (1)* provides the basic principles for the design and use of signs, signals, and pavement markings for all public roadways in the United States. The manual sets forth the warrants and standards as adopted by the Federal Highway Administration (FHWA). (It is understood by the authors that the MUTCD contains warrants, standards, description and guidance; however, throughout this paper all material in the MUTCD will be referred to as "standards.")

The requirements for the size, shape, and placement of various traffic control devices have been developed over the years. The American Association of State Highway and Transportation Officials (AASHTO) published a manual for uniform standards for rural highways in 1927. The National Conference on Street and Highway Safety published a manual for urban streets in 1929. FHWA and AASHTO formed a joint committee (NJC) and published the first MUTCD in 1935. Subsequent revisions to, or editions of, the MUTCD were published in 1939, 1942, 1948, 1954, 1961, 1971, and 1978.

As the MUTCD has evolved through the years, changes have often been made on a piecemeal basis. Some portions of the manual have changed very little since the earliest editions. Many of these older standards probably stem from subjective judgments made 40 or more years ago. Of those standards that have changed, some undoubtedly have an objective basis. In many cases, however, the basis for a change in the manual is obscure; documentation is lacking, and it is likely that many changes were made as the result of collective subjective opinion by the groups responsible for the continuing development of the manual.

The foregoing observations suggest that many of the basic elements or standards in today's MUTCD may not adequately serve the needs of the 1980s. Some of the basic standards that have been accepted as gospel may be deficient. For example, the 3.75-ft driver eye-height standard for marking no-passing zones was accepted for more than 20 years. Then, suddenly, the traffic engineering community realized that it was no longer adequate because of changes in vehicle design, and a value of 3.50 ft was adopted. Undoubtedly, other standards embedded in the manual are also obsolete. The basic objective of the research reported herein was to identify those standards so that needed traffic control device research can be programmed and conducted.

Identification of standards that may be obsolete is a difficult task because no single comprehensive source of historical and technical information exists to document the reason changes to the manual were made. To overcome this obstacle and to provide a comprehensive source of information for the MUTCD for future use, a comprehensive computerized filing system, which documents historical changes to the manual and relevant traffic control device research, was developed.

DEVELOPING RESEARCH NEEDS FOR MUTCD STANDARDS

The research review process combined a committee review, expert screening of standards, a computerized search of abstracts, establishment of research priorities, a library search for the high priority articles, an empirical evaluation of research reviewed, and a listing of standards by need for additional research. Figure 1 shows this process in more detail.

In the initial step of the MUTCD evaluation the basic standards included in the document were identified. "Basic standards" exclude very general statements that reflect a broad attitude but do not provide specific guidance. For instance, Sections II-A-8 and II-A-30 contain broad statements on sign standardization and maintenance that really cannot be construed as standards, but as a general philosophy of practice.

Therefore, the study team developed a list of basic standards

P. S. Shapiro, J. Loewen, and V. Siaurusaitis, COMSIS Corporation, 11501 Georgia Ave., Wheaton, Md. 20902. J. E. Upchurch, Department of Civil Engineering, Arizona State University, Tempe, Ariz. 85287.

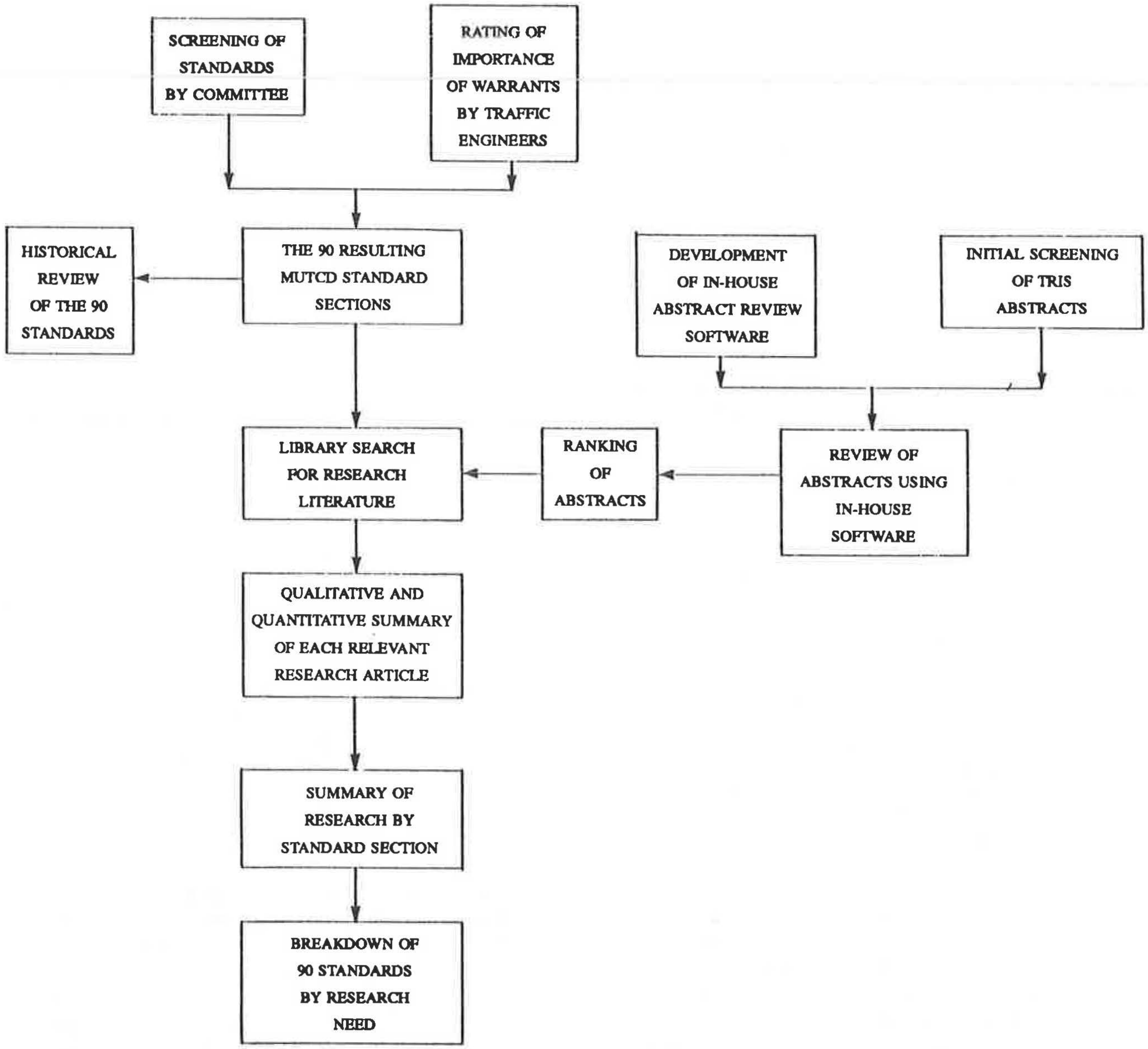


FIGURE 1 Research review process.

derived from review of the manual for each traffic control device, or set of devices. Standards for each device or set of devices were extracted from the manual, and for each device the warrants and standards were reported as follows:

- Physical attributes (size, shape, color, etc.);
- Message/meaning/legend (intended meaning, wording, or image content, lettering, etc.);
- Illumination/reflectorization;
- Warrants;
- Placement; and
- Other [depending on the type of devices, some other aspects may be important (e.g., mounting), or some standards may not readily fit one of the other categories].

The 1978 Manual section number and title was used as the key identifier with the substantive information from each standard placed under its respective heading. By "substantive" information, it is meant that only statements that were general, reflecting broad attitudes or general management approaches, were omitted. Thus the use of every device was fully specified by the list of warrants and standards in a uniform format without superfluous information.

Once the standards and warrants were reformatted as described previously, the research team screened them to determine the standards that would most likely need further research. This subjective screening process was based on the following three factors: importance of the device or standard, lack of a known research basis, and potential of finding research that describes the basis of a device or standard. As a result of this screening process, 517 standards and standard parts were identified as warranting further consideration. These were then compiled into a manual that was distributed to a nine-member panel of traffic engineers for further evaluation. This expert panel consisted of currently active traffic engineers working for operating agencies, FHWA staff from the Office of

Traffic Operations and Office of Research, experts with a long history of activity on the committees responsible for the 1978 Manual and its predecessors, and researchers in traffic operations.

The panel members were asked to independently assign a priority to the standards and standard parts according to a scale between 1 and 3 with 1 indicating low importance and 3 a higher priority for further investigation. The panel was also asked to suggest any additional areas of the manual that should be addressed.

The rating by individual panel members for each of the 517 standards and standard parts was compiled, and a frequency distribution of scores was prepared. All standards or standard parts that received a score of 15 or greater were selected for further review.

Following the ratings, the panel met to discuss the results of the prioritization and the issues involved. Some standards identified by the panel but not included in those selected by the research team were added to those to be reviewed further. The result was a list of 90 standards to be subject to further scrutiny of their research basis. These were then grouped into four traffic control device categories. The categories, followed by the number of standards in each category, were as follows: signs (44), pavement markings (13), signals (10), and construction and school zones (23).

RESEARCH REVIEW PROCESS

Abstract Search

The use of a microcomputer search system to find research for the selected standards expedited the research process. Relevant research for the selected standards was located using the Transportation Research Information Service (TRIS) data file. More than 6,000 TRIS abstracts were identified and transferred to nine-track computer tapes and then downloaded to a PC hard disk. Two microcomputer application programs developed for this project, KEYWORD and XREF, facilitated the abstract search.

Computerized Search Process

The process by which abstracts were located for a selected standard involved three steps. The first step required a review of the wording for the selected standards to extract keywords to be used by the KEYWORD program. Next, the XREF program was run for two or three keywords to generate a more refined listing of abstracts relevant to a selected standard. Finally, abstracts generated by the XREF program were reviewed by the research staff which determined whether or not the reports would provide some research basis for the selected standards.

The KEYWORD program created a file of important words used in each of the manual sections. For example, to find research relating to stop signs the user would enter KEYWORD STOP. A file would then be created of all abstracts, identified by a unique number, which contained the word STOP.

The XREF program allowed for more specific abstract searches. This software cross-referenced two or three words from the keyword file. For example, if the users were interested in finding research related to stop sign location, they would type in XREF STOP LOCATION. This would instigate a two-step interaction. First, all of the abstracts containing the term STOP would be flagged. Of these abstracts, the ones containing the word LOCATION would then be found. Thus a file of abstracts containing both the desired terms would be created.

PRIORITIZATION PROCESS

Before starting the literature review, which was conducted at the U.S. Department of Transportation library and other libraries, abstracts were ranked by their relevance to the question being asked of a particular standard. Ranking by priority was very important because it would be impractical to review all the abstracts identified for each standard.

Prioritization, ranking, and review of relevant research for the selected standards was conducted in a two-tier fashion. Prioritization and ranking processes and forms were developed by the staff. A review of 25 articles by each of 3 members of the staff at the beginning of the review process indicated the consistency of staff rating. Comparing individual staff ratings for the 25 articles, it was found that ratings were consistently uniform for the 3 staff reviewers. Because of the large number of research articles, it was not feasible for each staff member to review each article. Judging from the success of the review of the 25 articles, it was decided that they would be divided among the research staff for review.

Abstract Review Process

Abstracts that were believed to be of greater explanatory power for a standard were reviewed first. The ranking system developed for this step of the research process was as follows:

- A = research based directly on or closely related to the question(s) under examination;
- B = tangential research issues related to the question(s) (but not directly addressing the issue);
- C = discussions related to the question, but not based on research; and
- D = not relevant (delete).

The TRIS search identified 5,893 abstracts related to the 90 standards; 3,288 were referenced as potentially useful. The staff then flagged 1,314 of these abstracts and ranked 371 as A, 250 as B, 308 as C, and 385 to be deleted. Because of the limited amount of time available to locate research related to the abstracts, the researchers concentrated on locating the research described in the 371 abstracts ranked as having the greatest amount of relevance to the selected standards, ranked A. The results of this review are given in Table 1.

Library Search

The next step in the research review process was to develop a way to review A abstracts quickly and efficiently to determine if they contained research that supported the standard wording.

TABLE 1 MUTCD ABSTRACT SEARCH

Section	Standards/ Standard Parts	Total No. of Abstracts Identified Through TRIS Search	Abstracts Identified by XREF Search	Abstracts Initially Reviewed	Detailed Review				
							Rank Order		
					Detailed	Kept	A	B	C
Signs (II)	44	2,394	1,755	695	194	501	166	119	216
Markings (III)	13	1,940	532	171	60	111	46	30	35
Signals (IV)	10	890	565	246	74	172	76	59	37
Construction and school (VI, VII)	23	669	436	202	57	145	83	42	20
Total	90	5,893	3,288	1,314	385	929	371	250	308

Research Form

The research staff developed an MUTCD research form for this review. The form contained items that systematically determined the adequacy of the research. The type of research, geographical location, date, objectives, conclusions, assumptions and biases, methodology, sample size, quality, and significant findings were included in the evaluation of each research effort.

A form was completed for each research report reviewed. A completed form indicated the study team's impression on the quality of the research as well as any significant findings relevant to the standards. At the end of each research form, the staff provided a final ranking on the quality of the research, as follows:

1 = research findings completely address standard wording and fully answer question(s) posed.

2 = research findings partially address standard wording and partially answer question(s) posed.

3 = research findings dispute standard wording and do not answer question(s) posed.

Summary Form

Once all the "A" articles had been reviewed and an MUTCD research form completed for each one, an overview of the research found for each individual standard was conducted. A summary form outlined all the research that had been located for a standard, compared the research to the standard wording, and posed questions on the adequacy of the research and whether or not a research study should be designed. Completed forms and research articles were then reviewed by a member of the team familiar with available research. His familiarity with current research would supplement research not found during the review process.

Quantification Form

The final step in the research process was to determine which standards required additional research. By reviewing the research adequacy forms, the staff could determine the standards that lacked research. Taking the process one step further, a quantification of research adequacy by standard form was developed to rank certain criteria for each standard on a scale from 0 (none) to 5 (very good). This form indicated the type of research performed as well as the quality of the research criteria for each standard. Figure 2 shows the form used in this step.

RESEARCH FINDINGS

The research review process, conducted in a systematic fashion, suggested where further research was particularly needed.

MUTCD standards with little or no apparent scientific research verification as well as standards with significant relevant research were noted. In addition to the literature search, TRB circulars and FHWA documents were consulted to determine current research efforts. Further insight into the status of MUTCD-related research was gained through discussion with transportation professionals in government and the private sector. On the basis of the preceding findings, the research team made judgments about the need for additional research for each standard.

Many traffic control devices and warrants are likely to benefit from further evaluation, improved design, or a better understanding of driver capabilities and behavior. The process was not devised to examine the possible research that might be

TABLE 2 SECTIONS OF THE MUTCD IDENTIFIED TO HAVE A SIGNIFICANT NEED FOR ADDITIONAL RESEARCH

Control Category	Number	Section Title
Signs	II-B-5	Stop sign
	II-B-6	Multiway stop signs
	II-B-8	Warrants for yield sign
	II-B-9	Location of stop sign
	II-B-32	Placement of urban parking signs
	II-C-5	Curved sign
	II-C-21	Narrow bridge sign
	II-D-5	Lettering style
	II-D-6	Size of lettering
	II-E-6	Reflectorization or illumination
	II-E-26	Advance guide signs
	II-F-13	Color, reflectorization, and illumination
Markings	II-B-1	Centerlines
Signals	IV-B-10	Illumination of lenses
	IV-C-2	Warrants for traffic signal installation
Construction and maintenance	VI-C-2	Channelization
	VI-G-3	Signs

STANDARD								
FIELD OBSERVATION								
SIMULATION MODEL								
DRIVING SIMULATOR								
LABORATORY								
STATISTICAL ANALYSIS								
SAMPLE SIZE								
SIGNIFICANT FINDINGS								
QUALITY OF RESEARCH								
AGE-IS RESEARCH CURRENT?								
NEED FOR ADDITIONAL RESEARCH								

RATING SCALE	NEED FOR ADDITIONAL RESEARCH SCALE
5 - very good	5 - high need for additional research
4 - good	4 - medium need for additional research
3 - average	1 - low need for additional research
2 - poor	
1 - very poor	
0 - none	

FIGURE 2 Quantification of research adequacy by standards.

directed at each MUTCD standard; rather, it was to identify those standards, or underlying issues, for which no adequate technical basis appears to support the manual requirement.

Signs

Of the four major control categories evaluated, traffic signs proved to be the most prominent. Not only were there more of these standards identified than for other categories, but there were more of these identified that had a significant need for additional research (see Table 2). This may be because signs are the oldest devices. Thus, a large amount of research is outdated. In addition, many of the new sign standards have been added without corresponding research.

Notable among the signs flagged for additional research are the intersection controls. Stop and yield signs—particularly the former—accounted for one-third of the "significant need" sign standards. This is a result of disagreement in the research as well as a lack of research addressing specific questions because a substantial amount of research was conducted for both devices.

Also notable among the various issues related to signs are visibility standards. Two aspects of this issue are notably in need of further research: (a) warrants related to reflectorization and illumination that occur in Sections II-E-6 and II-F-13, and (b) standards related to lettering style and size that occur in Sections II-D-5 and II-D-6. The issues of sign lighting and legibility are also included in another matter needing further research, one that is becoming increasingly important—the needs of the aging driver.

The demographics of the United States are changing; the average age of motorists is increasing. With the growing number of drivers over 50 years of age, visibility and legibility are vital issues. Visual acuity and other visual capabilities of drivers decrease with age. Perception and reaction time also increase. Thus the basic premises of standard design, particularly sight distance and perception reaction time, may need to be readjusted.

Pavement Markings

Among the pavement markings standards, only the issue of centerline markings was believed to be significantly in need of additional research. This issue is related to the need for better visibility, particularly in relation to wet-night driving. There are also serious concerns about the driving public's understanding of lane markings.

Signals

Signals were, in general, well researched. None of the standards related to this category of device was without research.

Of the two selected standards in the "significant need" category, the warrants for one have changed since 1978. Three new warrants for traffic signal installation (Section IV-C-2) became effective on January 1, 1985. As noted in the December 1985 *ITE Journal*, these warrants could lead to an increase in the number of signalized intersections throughout the country.

The primary research need associated with selected Standard IV-B-10 is daytime versus nighttime visibility. It appears that nighttime visibility is adequate for all three colors. Daytime visibility of traffic signals is often limited, particularly when subject to the direct rays of bright morning or afternoon sun. Green is the least identifiable of the three colors. Thus the visibility of green signals during these critical times is an important research issue.

Construction and Maintenance

One issue that arose throughout the study was the color of construction and maintenance signs. This topic cuts across category classification. It was first mentioned as part of Section II-A-11, the sign color section. Subsequently, it appeared as part of Section VI-B-1, Design of Signs. The question of whether or not orange is the appropriate color also appears in reference to Design and Application (Section VI-B-13), Cone Design (Section VI-C-3), and Drum Design (Section VI-C-6).

RESEARCH RECOMMENDATIONS

Several sections of the MUTCD that have a significant need for additional research were identified in the preceding section of this paper. The following recommendations identify eight high priority areas in which further research on traffic control devices is needed.

A number of important issues surfaced throughout the MUTCD evaluation of various standards. Many standards that are inadequately supported by research were identified in the review process. If there was a specific lack of research in more than one standard, in most cases it proved to be a major issue. The eight major MUTCD research issues are identified as follows:

- Shall, should, may.
- Symbols versus word.
- Yield versus stop.
- Construction and maintenance signs.
- Reflectorization and illumination.
- Compliance.
- Older drivers.
- Design drivers.

These eight issues are recommended to have the highest priority for future MUTCD research, an analysis of each issue is presented next.

Shall, Should, and May

Although the manual defines these terms (Section I-A-5), ambiguity remains and the basis for selection of a particular term for a given application is not apparent. Beyond these issues, two other closely related issues require research. First, how do local traffic authorities actually interpret or respond to each of these terms, and for what reason? It appears that a "should" ("advisory" condition) is often *de facto* a "shall" ("mandatory" condition) if a local jurisdiction is concerned with protection from potential tort liability suits. Second, given the actual response of local authorities to these terms, what is the implication of the use of one term versus another for the safety and operational efficiency of the traffic system?

The effects on overuse of traffic control devices, inappropriate applications, or failure to implement where needed, may be tied to the choice of terms. An appropriate study could examine the operational and safety effects, costs, and tort liability implications of the choice of the terms "shall," "should," or "may." The study may also develop guidelines for the choice of terms and review current MUTCD standards under these guidelines.

Symbols Versus Word Legends

The use of symbol signs versus word legends was an important issue. The question was whether or not word legends can be replaced by symbols and still be clearly and rapidly understood by the motorist. The incorrect interpretation of these symbol signs poses a serious danger to both motorists and pedestrians.

A substantial amount of research has been conducted on this topic. The MUTCD general philosophy has evolved to reflect the strong international trend toward greater use of symbols. In general, symbol signs have been found to be more effective than word messages in terms of their perception time and legibility distance. One study found symbol signs identified at more than five times the distance of signs with word legends. Another research article indicated that symbol legibility can be considerably increased by improving symbol design. Other findings indicated that comprehension of symbols is reduced by the addition of information such as words or prohibitory symbol elements.

Despite the amount of research on symbol versus word signs, the question of superiority remains unresolved. Symbols generally perform better, but this has not been universally true. Many comprehension problems exist for current MUTCD symbol signs.

The authors review suggests that an important general question is, Under what conditions are symbol signs preferable to word signs? To improve symbol sign design, there may be a need for standards that parallel existing standards and guides for word legends. Accepted graphics principles analogous to letter height or stroke width could be useful. There are no criteria for developing, selecting, or evaluating a symbol for comprehension, legibility, and so forth. Research is needed to determine general criteria for symbol sign use and design: when to use a symbol versus a word legend, how to determine the pictorial content, and what the design and evaluation criteria are. MUTCD symbol signs should be reviewed against these considerations.

Yield Versus Stop Signs

Research indicated that yield signs are underused. In many cases a stop sign could be replaced by a yield sign with no adverse impact on driver safety or the efficiency of road use. Standards affected by this issue include: Section II-B-5: Stop Sign, Section II-B-6: Multiway Stop Signs, and Section II-B-8: Warrants for Yield Signs.

Color of Construction and Maintenance Signs

The color of construction and maintenance signs was a major issue. The major question was whether orange was the best color for use on construction and maintenance signs.

Controversy continues regarding the adequacy of orange signs (with black legend) in terms of both perception and comprehension. The relative visibility and legibility of orange signs has been questioned. Perhaps equally important, but less researched, is the conspicuity of orange signs for unalerted drivers in realistic settings. On the comprehension side, some evidence suggests that the general public does not adequately understand the color coding. Again, an important aspect has not received adequate attention: How well does the orange sign convey the sense of hazard and the need to take some action? If orange were found to be less effective in terms of visibility or meaning, the logic behind independently color coding construction and maintenance signs should be reviewed. The performance of black-on-orange signs (photometrically, perceptually, and in meaning) should be evaluated (a) against objective performance criteria, such as required legibility distance; and (b) relative to the performance of alternative colors.

A recommendation to continue, modify, or drop the construction and maintenance color coding should be made with explicit reference to the data and logic involved.

Reflectorization and Illumination of Traffic Control Devices

The issue of reflectorization and illumination was important with respect to both signs and markings. For signs, reflectorization is a factor in installation of overhead signs, visibility of street name signs, and the importance of sign colors appearing essentially the same by night and day. Illumination was a key issue because with the changing technology for reflective material and the increasing cost of electric power, signs that are illuminated might be replaced by reflectorized signs. For pavement markings, reflectorization is a factor in the visibility of longitudinal pavement markings, object markers, and raised pavement markers.

Most MUTCD standards and recommendations concerning reflectorization, illumination, and options among these are quite vague. A study to develop performance criteria for such standards may lead to greater uniformity and ensure adequacy.

Lack of Compliance with Traffic Control Devices

Many enforcement agencies and highway authorities believe a serious and growing problem is motorist noncompliance with traffic control devices. Although no objective data were encountered to confirm that the problem is increasing, the literature contains a variety of studies that evaluate noncompliance with specific traffic control devices.

A number of compliance-related issues require evaluation. These include the magnitude of the noncompliance problem, its safety implications, and identification of those traffic control devices that constitute the greatest problem. The noncompliance problem may be improved by better MUTCD standards for design or placement of devices or both. Research should analyze why compliance problems occur with particular devices and identify improvements in traffic control device design.

The Needs of the Older Driver

The aging of the driver population as U.S. demographics change has become a concern. Research supporting many MUTCD standards has not incorporated evaluation of the capabilities of older drivers. Many age-related changes in ability and behavior, both visual and nonvisual, may influence the adequacy of warrants and devices. There is a need for a comprehensive review and evaluation of the needs of older drivers and the adequacy of current standards. (This is part of the larger "Design Driver" issue, discussed later.)

The importance of evaluating age-related problems comprehensively has been emphasized in this review. Most attention thus far has focused on visual decrements such as acuity, glare sensitivity, and so forth. This has obvious implications for letter size, reflectivity, and illumination. However, other age-related decrements, in factors such as speed of information processing or ability to time-share simultaneous demands, pose equal demands on MUTCD standards. This affects numerous factors, including device location, information content (particularly guide signs), temporal aspects (e.g., duration of the yellow phase, clearance intervals, advance signing), and symbol comprehension (which has frequently been shown to be poorer for older groups).

Future research must include a comprehensive review of age-related changes, both visual and nonvisual, that affect MUTCD standards. The impact of current inadequacies on the older population should be evaluated in terms of safety, operational efficiency, and the discouragement of mobility.

Improved criteria to address older driver requirements should be developed and current devices evaluated under these guidelines.

The Design Driver

It is suggested that the manual add a section on design driver criteria. This would include factors such as eye height, acuity, and response time for various actions (recognition, braking, etc.). The factors could be broken down by percentile (50th, 85th, 95th), or key driver groups that may be of concern (by age, condition, etc.). Also, it should include the appropriate formulas for combining the basic characteristics to derive other key quantities, such as legibility distance, decision sight distance, and so forth.

Such a section would achieve several important goals. First, it would provide a clear set of consensus criteria for use in evaluating the adequacy of standards for devices. Second, it would allow the adequacy of devices to be periodically reviewed as changes occur in the vehicle fleet, roadway features, or the knowledge and assumptions about driver performance. Third, it would permit well-defined performance-based standards. It would permit standards such as, "the sign should have a minimum decision sight distance of X," rather than specifying some single size to cover possibly quite different situations.

FILE MANAGEMENT SYSTEM

In addition to identifying areas in which the need for additional research was great, a second goal of this study was to make the MUTCD accessible for review through a computerized comprehensive file management system. To maintain and update the information that was compiled for the MUTCD evaluation, a file management system was developed. The filing system contains (a) all the information obtained during the research of the selected standards, (b) a historical review of the selected standards, and (c) the 1978 MUTCD and requests for changes. The system was developed using a microcomputer that stores all the MUTCD information and allows the user to make changes to the data base.

The MUTCD File Management System (FMS) is a menu-driven search and maintenance program created in dBASE III (registered and copyrighted by Ashton-Tate). The system allows the user to access any of the created files using the standard identification number or a specific keyword. A unique numbering system to find information pertaining to a given standard facilitates searches. The numbering system in the 1978 MUTCD was reformatted to a system that would work in the FMS. For example, MUTCD standards II-B-5, III-A-3, and

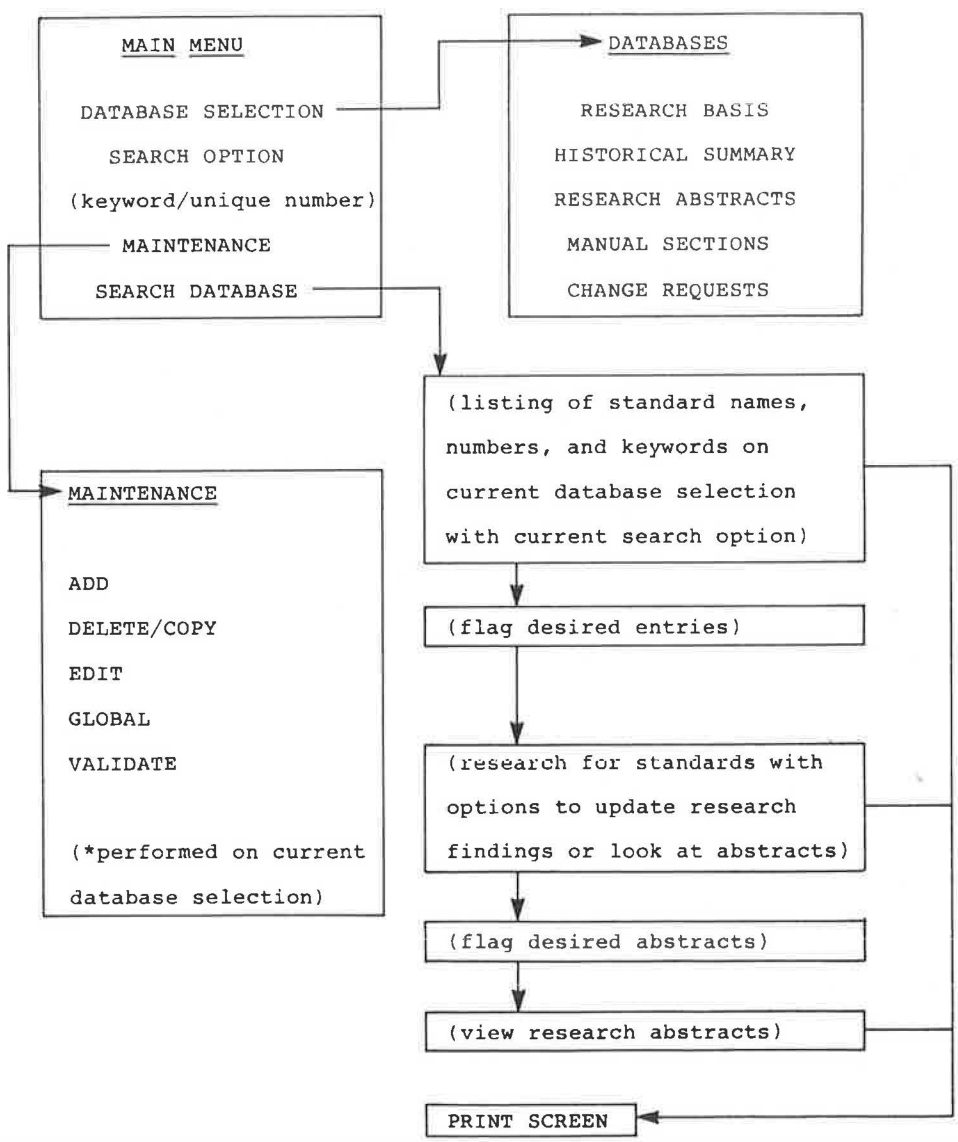


FIGURE 3 File management system flow chart.

VI-D-1 would be 2B#5, 3A#3, and 6D#1, respectively, in the FMS. The system is set up so that the user can locate information for a particular standard, search for information pertaining to a particular topic, update the data base, or add new information to the data base. The following is a description of the functions and screens available in the FMS for MUTCD retrieval or maintenance, or both. Figure 3 is a flow chart of the FMS that helps the reader understand the hierarchy of screen retrieval and access.

When the user logs on to the FMS, the main menu appears and prompts the user for the type of function to perform. The user can specify four options at this point:

- *Data base selection*—prompts the user for the data base to be manipulated (i.e., manual sections, research basis, change requests, historical summary, research abstracts).
- *Search option*—determines if data base manipulation will be performed using the unique numbering system or by keywords related to a given standard.
- *Maintenance*—performs various manipulations on the data base selected.
- *Search data base*—performs a search on the data base selected using the search option selected.

The data base selection and search option set up pointers within the FMS so that the maintenance and search data base options can be performed. In addition to these four options, at this point the user can exit the FMS.

Three of the five data bases available for manipulation only contain information pertaining to the 90 standards addressed during this MUTCD standards evaluation. These data bases are research basis, historical summary, and research abstracts. Research basis is a review of located research that has been completed for the 90 standards. Historical summary is a manual-by-manual history of the 90 standards starting with the 1927 MUTCD where applicable. Research abstracts are the pertinent research reports found for the 90 selected standards. A typical abstract contains the title, author, and location of a research report followed by a brief description of the major points in the research. These three data bases can be added to as research is completed for those standards outside of the 90 already evaluated. This is performed in the maintenance option.

The manual section data base is the 1978 MUTCD. The change requests data base is a summary of all proposed changes to the 1978 MUTCD. Modification to these data bases is also performed in the maintenance option.

The maintenance option provides the user with five functions to manipulate the FMS. The user can change or delete existing information or add new information to the FMS. Specific functions for this option include

- Add. Adds new information to current data base selection.
- Delete/copy. Deletes existing information from the selected data base, or copies that information onto a work file outside of the FMS.
- Edit. Edits the keyword or subject fields for existing standards; these fields are referenced when the keyword search option is specified at the main menu.
- Global. Edits all records with missing keyword or subject fields.
- Validate. Updates the master file with any changes the user may have made during the maintenance option.

The search data base option will display all information in the FMS in the selected data base using the search option specified. For example, if the user performed a keyword search on the word "stop," the FMS will display all standards with the word stop used as a reference in the keyword or subject fields. At this point the user can flag individual standards and view the research that has been completed for that standard as well as the abstracts pertaining to the standard. New or additional research findings can be entered into the search data base option.

The MUTCD File Management System is an easy system to use because of its convenient menu-driven format. Its value is maximized if it continually is used to update changes that occur to the existing 1978 MUTCD. The groundwork has been laid to facilitate the easy access of research and information related to the MUTCD. If the FMS is completed for the entire manual, it would provide a comprehensive reference for information pertaining to uniformity of traffic control devices.

CONCLUSION

Described in this paper is a project whose primary objectives were to locate areas of research need and establish a file management system for the *Manual on Uniform Traffic Control Devices (1)*.

Admittedly, the basis for most of the standards was not located. Whether developed through research, in committee meetings, or over lunch, the origin of most of the standards, with several exceptions, is not certain. The research adequacy of selected MUTCD standards nevertheless has been investigated and reviewed. An evaluation of the traffic control device warrants needing further study has been completed and recommendations have been made. In addition, the MUTCD has been systematically placed in a data base management software package to facilitate review.

The review of research has highlighted areas in which future studies are needed. Significant among these is the need for examination of the aging driver and design driver issues.

With its manual, research, historical, change, requests, and abstract components, the PC-based file management software allows the user to easily locate information related to traffic control devices. With the keyword and standard search options, key terms and warrants are found with a few simple keyboard strokes.

Of equal significance, this project has provided a model for future MUTCD studies. The groundwork has been laid to facilitate subsequent investigations related to the manual and to allow for easy access of related research.

Of critical importance is that the process developed to identify research needs regarding the MUTCD and the file management system developed to organize information are both used. It is important that all research related to traffic control devices completed from this time forward be added to the FMS so that, in the future, the bases of individual standards can be established. Furthermore, the basic research needs identified, such as the development of design driver criteria, should be addressed in future research and future modifications to the MUTCD.

Finally, it was impossible to research the basis of all MUTCD standards; therefore, further efforts should be made to determine the basis of standards not addressed by this research.

ACKNOWLEDGMENT

This research was sponsored by the Federal Highway Administration, U.S. Department of Transportation. More detailed documentation on this research project may be found in the project final report entitled "Evaluation of MUTCD Selected Standards" published by the COMSIS Corporation.

REFERENCE

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

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 Federal Highway Administration. This paper does not constitute a standard, specification, or regulation.

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

Evaluation of Wide Edgelines

J. W. HALL

In many states, single-vehicle run-off-the-road (ROR) accidents constitute one of the most significant traffic accident problems. Described in this paper is a study of the effectiveness of one potentially useful countermeasure—the application of 8-in. wide edgelines. The critical rate technique was used to identify approximately 530 mi of rural two-lane highway with unusually high ROR accident rates. In 1984, 100 of these miles were treated with wide edgelines, and the following year, an additional 76 mi were marked. The remaining mileage was used for comparison purposes. The accident experience on the treatment and comparison sections was monitored after the application of this countermeasure. From the research it is concluded that wide edgelines do not have a significant effect on the incidence of ROR accidents. In addition, this treatment does not have a significant effect on the rate of ROR accidents at night or on curves, or on accidents involving the opposing flow of traffic. It is recommended that this treatment be discontinued on rural highways in New Mexico.

On a nationwide basis, single-vehicle run-off-the-road (ROR) accidents account for approximately 38 percent of all highway fatalities *(1)*. The two predominant collision types within this set of accidents are overturning and impacts with fixed objects. Total statistics for New Mexico are similar, with single-vehicle ROR accidents responsible for 41 percent of the highway fatalities *(2)*; however, because of the relatively clear roadsides in the state, a greater proportion of the ROR accidents involve overturning *(3)*. Clearly, the consequences of a vehicle departing from the traveled way are a function of roadside characteristics, specifically the presence of obstacles and the nature of roadside slopes.

For the past two decades, the technical literature and federal standards have promoted the use of forgiving roadside designs *(4, 5)*. These designs, characterized by flat side slopes, removal of unnecessary fixed objects, and the use of attenuators and breakaway supports, have been used extensively on freeways and some rural highways. In response to the increased emphasis on highway safety in the 1960s, annual highway fatalities have decreased by nearly 20 percent, and the fatality rate per 100 MVM has dropped by 50 percent *(6)*. These dramatic improvements can be attributed to numerous programs affecting the highway, the vehicle, and the road user. During this same time period, fatal ROR accidents have also decreased. However, the improvement in this area has not been as great as that cited for all fatal accidents noted above; and, in fact, there is evidence that after all the effort devoted to clear roadsides, fatal ROR accidents now constitute a larger share of all fatal accidents than they did 20 years ago *(1)*.

During the past decade, a number of studies have suggested some reasons for this unexpected result. There is, for example, a growing body of knowledge indicating that certain fixed objects formerly believed to be "safe" are actually hazardous under some impact conditions *(7, 8)*. Occupants of small vehicles are vulnerable in impacts with certain barriers and with breakaway objects that post little problem for occupants of larger size passenger cars *(9)*. In addition, there is evidence that the 3:1 front slope criteria often cited as the warrant for longitudinal barrier installation may be too steep *(10)*. Collectively, the results of these studies indicate there is still a lot to be learned about roadside safety.

A second set of studies has established that roadway geometrics contribute to ROR accidents *(11, 12)*. Specifically, locations with adverse alignment, including sharp curves to the left, downgrades, and inadequate superelevation, all have unusually high accident experience. Although this result is not particularly surprising, the concept of using adverse roadway design characteristics as criteria for selecting sites for roadside safety improvements has not become widely accepted. As a result, agencies may sometimes devote resources to providing clear, flat roadsides at locations where vehicle departure from the road is comparatively unlikely.

There are basically two approaches to reducing the problem of run-off-the-road accidents. One approach is to provide a safe, traversable roadside that permits an errant motorist to regain control of the vehicle. This approach has the effect of reducing the severity of these incidents. Logic would suggest that, if all other factors are equal, the nature of the roadside should not affect the frequency of roadside encroachments, although one recent study suggests this may not be true *(13)*. An alternative approach is to improve the roadway to reduce the incidence of vehicle encroachment. Roadway realignment, shoulder widening, and the removal of edge dropoffs are some potential improvements in this regard. This approach may potentially reduce the frequency of these incidents. The relative cost-effectiveness of these two approaches obviously depends on the physical characteristics of the particular site.

The engineering community is committed to providing safe and forgiving highway designs. However, the cost of implementing roadway or roadside improvements on the extensive system of existing highways is an expensive proposition. In response to this situation, most states have developed schemes for assigning a priority to locations for improvement *(6)*. These techniques, which typically rely on previous accident experience, can help optimize the expenditure of limited funds available for remedial action. In many but not all cases the construction of new facilities can incorporate the appropriate safety features with little additional cost. As a practical matter, however, comparatively few miles of new highway are being constructed.

A number of states have experimented with other methods to

The University of New Mexico, Albuquerque, N.Mex. 87131.

reduce ROR accident frequency. A 1982 survey of state highway departments revealed a strong preference for the use of relatively inexpensive treatments such as chevron signs, delineators, and other traffic control devices *(14)*. These countermeasures can have a positive effect on a reasonably attentive driver, although they cannot eliminate all the factors that contribute to crash occurrence. Furthermore, this same survey revealed that few agencies had actually evaluated the effectiveness of these economical treatments.

Pavement markings constitute one class of countermeasures that has been evaluated in a number of studies *(15)*. The effectiveness of pavement markings appears to derive from several factors, including their ability to delineate a travel path, their placement on the roadway where the driver's attention is focused, and their relatively simple message. They also have several drawbacks, including deterioration as a result of traffic and environmental conditions, lower visibility on wet pavements, and blockage by snow and dirt. The commonly used treatments with relevance to ROR accidents are centerlines and edgelines. Although the results of studies of the effectiveness of pavement markings are not entirely consistent, the available data suggest that they have a small but positive effect on driver behavior.

The basic provisions of the *Manual on Uniform Traffic Control Devices* (MUTCD) with regard to markings on main rural highways *(16)* are fairly straightforward. Centerlines are yellow, edgelines are white, and both must be reflectorized. Normal line widths are 4 to 6 in. Rural two-lane highways with adequate width and speeds greater than 35 mph should have centerlines, whereas the application of edgelines under these conditions is at the discretion of the engineer. The desirability of edgelines is indicated by the requirement for their installation on rural, multilane divided highways, including Interstate highways.

In recent years, several sources have reported that wider edgelines, typically 8 in. can have an even greater benefit. The installation of these edgelines is certainly consistent with the provisions of the MUTCD. One study *(17)* found that wide edgelines, as opposed to 4-in. edgelines or no edgelines, caused drivers to assume a more central position in their lane and reduced the incidence of both centerline and edgeline encroachments. The researchers observed this improvement for normal drivers as well as for a set of drivers impaired with blood alcohol levels of 0.05 to 0.08 percent. Although these results are certainly encouraging, their relationship to accident experience has not been established. The data suggest that the use of wide edgelines reduces the potential for both ROR and sideswipe accidents, but the actual verification of this hypothesis requires an alternative study design.

In a previous study *(18)*, a procedure for the identification of roadway sections with unusually high ROR accidents was developed and applied to New Mexico's rural, non-Interstate highways. This procedure, described in more detail in the next section, led to the selection of sites warranting further study for possible remedial action. In light of the previous discussion, the application of wide edgelines was a good candidate treatment. The New Mexico State Highway Department painted 19 sections of road, a total length of 100 mi, in June 1984. The following year, New Mexico's informal program was expanded and modified to make it part of an FHWA study.

STUDY PROCEDURE

As discussed earlier, there is some evidence that the use of 8-in. edgelines, as opposed to 4-in. edgelines or no edgelines, can have a positive effect on vehicle tracking patterns. Previous research *(17)* revealed that these changes were near the borderline of statistical significance. The question of greater importance to traffic engineers is whether these minor alterations in driver behavior will produce a significant change in the associated accident experience. The study plan developed in this research to resolve this issue consisted of the following steps:

1. Identify sections of road with high ROR accident experience.
2. Paint 8-in. edgelines on some of the sections identified in Step 1 while the remainder were painted normally and used as comparison sites.
3. Monitor the *after* accident experience of the treatment and comparison sites.
4. Conduct appropriate statistical analyses to determine if any significant reduction in accidents occurred because of the treatment.

Site Identification

It must be noted from the outset that this project was attempting to determine if wide edgelines reduce accident frequency by a significant (in the statistical sense) and meaningful amount. Clearly this treatment would not have an effect on the severity of the accidents that do occur. It is, of course, quite possible that the treatment has no effect, and it is even conceivable that the net effect is detrimental. For example, the wide edgelines could cause motorists to drive closer to the centerline, thus increasing the incidence of opposite-direction sideswipe accidents. In any case, the best opportunity to examine their effectiveness is on sections of road that are experiencing unusually high rates of ROR accidents.

Other researchers *(19)* have described the use of the rate-quality control technique for identifying abnormal roadway sections. This technique compares the accident rates on individual sections of roadway to the systemwide average, and detects sections with rates that are significantly above statistically expected values. The approach also considers various levels of exposure on the different sections. The formula for calculating a road section's critical ROR accident rate (RC) at the 5 percent level of significance is given by

$$RC = RA + 1.645\sqrt{RA/m} + 0.5/m$$

where RA is the systemwide accident rate and m is the vehicle miles of travel on the particular section.

The critical rate is obviously greater than the systemwide accident rate. It decreases with increasing travel on the individual study sections. If the travel and the ROR accident experience on a section are known, the actual section rate can be calculated and compared with its critical rate. Within the limitations imposed by the quality of the traffic accident and travel data, sections on which the actual rate exceeds the calculated critical rate are said to be hazardous at the 5 percent level of significance.

This identification process can be implemented in a fairly straightforward manner once a few details are clarified. The selection of roadway sections for analysis is obviously a basic need. In cooperation with the New Mexico State Highway Department (NMSHD), it was decided to limit the study to rural non-Interstate portions of the federal-aid primary (FAP) and secondary (FAS) systems. The highway department maintains a roadway inventory, which subdivides these roadways into sections of variable length based on the original construction contracts, certain jurisdiction boundaries, major intersections, and other selected physical features. These sections vary in length from several hundred feet to 20 mi, but are reasonably homogeneous and have a constant design speed through each section. The inventory includes information on the length and average daily traffic (ADT) of the sections in a format suitable for computer processing.

The initial step in the analysis was to calculate the average ROR accident rate for the rural FAP and FAS road systems. The average accident rates, given by the sum of all rural ROR accidents for the 3-year period 1981 to 1983 divided by the travel on these systems, were 0.58/MVM and 0.96/MVM for the primary and secondary systems, respectively. Other characteristics of these roadway systems are given in the following table:

	FAP	*FAS*
System length (mi)	3,280	3,776
ADT range vehicles per day (vpd)	120–19,800	7–12,200
Average ADT (vpd)	2,890	1,110
Daily travel (mvm)	6.65	2.58
ROR accidents per mile per year	0.43	0.24

It was previously mentioned that New Mexico has a high incidence of single-vehicle ROR accidents. However, when 1-mi-long rural segments of these systems are examined for a 1-year period, 75 percent do not experience an accident of this type. Even during the 3-year study period 1981 to 1983, one-half of the 1-mi-long segments did not experience a single-vehicle ROR accident. In other words, on a statewide basis these events are relatively common, but their occurrence on individual short sections is infrequent.

The second step in the analysis involved the selection of roadway sections for use in the study. A computer program was developed to process the roadway inventory data and to combine adjacent sections with similar design characteristics and traffic volumes. When the traffic volume on one section was within 100 vpd of the volume on the following section, the two were combined and the vehicle miles of travel on the new, longer section was calculated. This process yielded 933 separate roadway sections (494 on the FAP and 439 on the FAS). The individual sections created by this technique ranged in length from less than 1 mi to more than 30 mi.

The next step in the analysis was the development of a computer program for the calculation of the critical rate on each of these sections. The program searched the 1981 to 1983 files for single-vehicle ROR accidents that occurred on the 933 rural sections and kept a running account of the number of these accidents for each section. The data were then merged with the inventory information to calculate the actual and critical rates for each section. In the case of the FAP system, the critical rate is given by

$$RC = 0.58 + 1.645\sqrt{0.58/m} + 0.5/m$$

This calculation was performed for each of the 494 FAP sections by using the appropriate vehicle miles of travel (m) for the 3-year period. The calculated critical rate was compared with the actual rate on each section, and the program identified 61 sections where the observed ROR accident rate for the study period exceeded the critical rate. Similar calculations for the FAS, using $RA = 0.96$, identified 89 sections on this system that were critical at the 5 percent level of significance. The finding that 150 sections (16 percent of all sections) were critical was unexpected. There are, however, two explanations for this situation:

1. These accidents are not uniformly distributed on the roadway system. As a result, during the 3-year study period many sections had no ROR accidents whereas others had substantially above average accident experience.
2. A number of the critical sections identified by this technique are quite short (<0.5 mi) where the occurrence of one or two accidents was sufficient to classify the section as critical. If problems exist on these sections, they could probably be more effectively treated with spot improvements rather than edgelines.

Critical Site Characteristics

The 150 sections identified in this process account for 15 percent of the vehicle miles of travel on rural FAP and FAS roadways and nearly 22 percent of the mileage on these systems. However, for this 3-year period, they experienced 37 percent of all the single-vehicle ROR accidents on these roadway systems. In other words, they are substantially more hazardous than typical sections of rural highway. As such, they may constitute good candidates for treatment with wide edgelines. This set of sites was reviewed with the NMSHD, and some additional criteria were established for the selection of the actual treatment sites. First, the sections selected for treatment were restricted to lengths of approximately 3 to 8 mi; this restriction was intended to facilitate the actual painting of 8-in. edgelines. Second, treatment sections were required to have at least 10 accidents during the preceding 3 years; this constraint eliminated a couple of apparently critical sections for which the miscoding of a single accident location would have changed the section from critical to noncritical. In addition, sections on multilane highways and sections where reconstruction activity was planned were dropped from the list of potential treatment sites.

The output from the application of the techniques outlined in the preceding paragraph was a set of 19 treatment sites (10 on the FAP and 9 on the FAS). The characteristics of these sites for 1981 to 1983 are summarized in the following table:

	FAP	FAS
Section length (mi)	54.71	46.41
Total ROR accidents	230	118
Daily travel (vm)	134,700	45,800
Accident rate	1.56	2.35
ROR accidents per mile per year	1.40	0.85

These approximately 101 mi of rural FAP and FAS clearly have single-vehicle ROR accident experience that is substantially above their systemwide averages. These 19 sections were painted with 8-in.-wide edgelines in June 1984. The other sections that were identified by the application of the critical technique constituted the comparison sites.

The original study plan called for monitoring the accident experience at the treatment and comparison sites for a period of 18 months, and then conducting an analysis to determine if a significant change had occurred at the treatment locations. The justification for using comparison sites is discussed in the original report *(20)* describing this study. The purpose of the comparison sites is twofold: to account for other changes in the highway transportation system that may contribute to crash reduction, and to account for regression-to-the-mean. In the more traditional but less reliable before-and-after study without comparison sites, any change in the accident experience after the site is treated is attributed to the engineering treatment. This simple approach overlooks the contribution of changes in other relevant factors in the highway transportation system, such as increased enforcement, new vehicle designs, changes in driver behavior, and so forth. In addition, several studies *(21)* describe the problem of regression-to-the-mean and its effect on accident studies. It can be demonstrated that, on the average, a group of sites with unusually high accident experience during one time period will tend to have lower accident experience in a subsequent period. Because locations are often chosen for treatment because of their high accident experience, it becomes difficult to separate the true effect of the treatment from the effect of regression-to-the-mean. However, with the use of comparison sites, which were also chosen because of their high accident experience, the analyst has a better opportunity to identify the true effect of the treatment. Specifically, the comparison sites would be expected to improve in the after period; thus the effectiveness of the wide edgeline should be a function of the difference between the changes at the treatment and comparison sites.

In the fall of 1984, the federal government expressed an interest in the effectiveness of this countermeasure. The NMSHD agreed to participate in the FHWA project by including data from its original study sites and by identifying and treating an additional 100 mi of roadway. The additional set of treatment sites were identified by using accident data from 1982 to 1984 and the critical rate technique. Some of these sections were in the set of comparison sites based on the 1981 to 1983 data, whereas others were not critical during the previous time period. This new set of treatment sites consisted of 14 sections (5 on the FAP and 9 on the FAS) with the mileage evenly divided between the two systems. The NMSHD field crews subsequently determined that two of these sections, with a total length of 24 mi, were not suitable for treatment with 8-in. edgelines; the remaining 76 mi were painted in July 1985.

By the end of 1985, there were three single-vehicle ROR accident data sets relevant to the study of wide edgelines:

1. Data for 19 sections of road with a total length of 101 mi marked in June 1984. This data set consisted of before data for a 41-month period before the treatment and a 17-month period after the treatment.
2. Data for 12 sections of road with a length of 76 mi. Eleven of these sections were painted in July 1985, and the remaining section was marked in October. Accident data were available for a 52-month before period, and with the one exception, for a 5-month after period.
3. Accident data for a set of comparison sites for the 5-year period 1981 to 1985. To facilitate the analysis, all accidents in June 1984 and July 1985, the two months that treatment sections were painted with wide edgelines, were dropped from the analysis. It was subsequently shown that the accident experience during these two months was virtually the same as the monthly average for the 5-year period.

In actuality, the traffic accident data for the preceding year do not become available on January 1 of the following year. It takes time for the accident reports to be assembled, checked and coded, and entered into the computer system. For a number of reasons, this process took a little longer than usual for the 1985 accident data, and a reasonably complete file was not available until April 1986. At this time, new computer programs were developed to compare the before and after accident data. The statistical testing mentioned in succeeding sections was conducted using contingency tables at the 5 percent significance level.

Analysis of All ROR Accidents

The first analysis evaluated the June 1984 treatment sites and the comparison sites, data sets 1 and 3 as previously identified. The results are given in the following table:

	FAP	FAS
Treatment Sites	10	9
Before ROR accidents	246	136
Before accident rate	1.37	2.25
After ROR accidents	74	69
After accident rate	0.99	2.74
Comparison Sites	16	22
Before ROR accidents	467	461
Before accident rate	1.22	2.26
After ROR accidents	150	175
After accident rate	0.95	2.07

The overall accident rate for both FAP and FAS treatment sites decreased from 1.59/MVM to 1.43/MVM, a decrease of 10 percent. During the same time, the accident rate at the FAP and FAS comparison sites dropped from 1.59/MVM to 1.34/MVM, a decrease of 16 percent. The latter decrease was expected due to the aforementioned principle of regression-to-the-mean. If the treatment were truly effective, it would be expected that the decrease at the treatment sites would be even larger. However, this is not the case for this set of study sites.

Only 5 months of after data are available for the sites painted in July 1985. The four FAP sites experienced an increase in the accident rate from 1.03/MVM to 1.26/MVM, a change that appears to be insignificant based on the small after sample size. The ROR accident rate at the eight FAS sites dropped from 1.51/MVM to 0.98/MVM. The overall ROR accident rate at these 12 sites decreased from 1.32/MVM in the before period to 1.09/MVM in the after period. During these same analysis periods, the comparison site rate dropped from 1.16/MVM to 0.96/MVM on the FAP, from 2.27/MVM to 1.56/MVM on the FAS, and from 1.54/MVM to 1.17/MVM for both systems combined. In other words, the overall accident rate at the treatment sites decreased by 17 percent while the rate at the comparison sites decreased by 24 percent. These numbers should be viewed with caution, however, because the treatment accident rates for the after period are based on very small samples (11 accidents on the FAP and 13 accidents on the FAS).

Nighttime and Curve Accidents

The tentative conclusion at this point in the analysis is that the sites treated with the wide edgelines did experience a reduction in single-vehicle ROR accident rates, but the reduction was less than that experienced at a similar set of hazardous comparison sites. It has also been hypothesized that the placement of wide edgelines may have an effect on certain types of accidents, specifically those occurring at night or on curves. To test this theory, data sets 1 and 3, as previously described, were subdivided into the following groups:

- Daytime versus nighttime accidents, and
- Accidents on straight roads versus curves.

Obviously, breaking the data sets into these categories further decreases the sample size available for the analysis. In addition, reliable exposure data do not exist for the amount of travel at night or on curves. Such values would clearly depend on the design characteristics of the road and the nature of the surrounding environment. These characteristics probably vary among the individual study sections, although there is no evidence to suggest that they differ systematically between the treatment and comparison sites. Because the intent of this analysis is to determine the relative effect of this treatment at night and on curves, it is not essential that precise travel figures be used in the rate calculations. Previous research *(10)* on single-vehicle overturning accidents in New Mexico suggests that for the rural highways examined in this study, about 20 percent of the travel may occur during the hours of darkness, whereas 15 percent occurs on curves. These factors were used in calculating the following single-vehicle ROR accident rates, but it is emphasized that the resultant values must be considered rough estimates.

The results of the nighttime and curve analyses are summarized in Table 1. At the treatment sites, 57 percent of the ROR accidents in the before period occurred at night, whereas in the after period, the value dropped to 48 percent. The corresponding figures for the comparison sites are 49 and 46 percent. The average nighttime accident rate decreased by 31 percent at the FAP treatment sites and by 41 percent at the FAP

TABLE 1 SINGLE VEHICLE ROR ACCIDENTS

	FAP	FAS	Both
Day versus Night			
Treatment, daytime	10	9	19
Before ROR accidents	110	54	164
Before accident rate	0.76	1.11	0.85
After ROR accidents	35	39	74
After accident rate	0.59	1.94	0.93
Comparison, daytime	16	22	38
Before ROR accidents	235	234	469
Before accident rate	0.77	1.44	1.00
After ROR accidents	93	84	177
After accident rate	0.73	1.24	0.91
Treatment, nighttime			
Before ROR accidents	136	82	218
Before accident rate	3.78	6.77	4.53
After ROR accidents	39	30	69
After accident rate	2.61	5.96	3.45
Comparison, nighttime			
Before ROR accidents	232	227	459
Before accident rate	3.04	5.57	3.92
After ROR accidents	57	91	148
After accident rate	1.80	5.38	3.04
Curve versus Straight			
Treatment, straight			
Before ROR accidents	101	39	140
Before accident rate	0.66	0.76	0.68
After ROR accidents	37	14	51
After accident rate	0.58	0.65	0.60
Comparison, straight			
Before ROR accidents	256	169	425
Before accident rate	0.79	0.98	0.85
After ROR accidents	91	56	146
After accident rate	0.68	0.78	0.71
Treatment, curve			
Before ROR accidents	145	97	242
Before accident rate	5.37	10.68	6.71
After ROR accidents	37	55	92
After accident rate	3.30	14.57	6.14
Comparison, curve			
Before ROR accidents	211	292	503
Before accident rate	3.69	9.56	5.73
After ROR accidents	59	119	178
After accident rate	2.48	9.37	4.88

comparison sites. The reductions for both types of sites were considerably smaller on the FAS system. Except for the FAS comparison sites, the reduction in nighttime accident rates is greater than the reduction in daytime accident rates. However, the data do not support the contention that wide edgelines produce a significant reduction in nighttime ROR accident rates.

The most striking characteristic of the analysis based on roadway curvature is the large percentage of single-vehicle ROR accidents that occur on curves. (With respect to this variable, it should be noted that the decision of whether a roadway is straight or curved is a judgment made by the investigating officer; it is quite possible that different officers might classify the same site differently.) During the before period, 63 percent of these accidents at the treatment sites and 54 percent at the comparison sites occurred on curves. At both types of sites, the percentages were about 8 percent higher on the FAS system. These percentages are higher than those found in previous studies *(3, 10)* in New Mexico. In the after period, the ROR accident rates decreased on curves on the FAP system (by 39 percent at the treatment sites and 33 percent at the comparison sites), while increasing by 36 percent at the FAS treatment sites. Overall, there was an 8 percent reduction in curve accident rates at the treatment sites, and a 15 percent reduction at the comparison sites.

Opposite-Direction Collisions

It has also been suggested that the application of wide edgelines may have an effect on the frequency of two-vehicle opposite-direction (OD) accidents. On the positive side, it could be argued that the treatment provides the driver with good guidance information that is especially valuable when he is partially blinded by the headlights of an oncoming vehicle. It is possible that drivers may shy away from a wide edgeline, thus moving them closer to vehicles traveling in the opposite direction. As a practical matter, both of these situations may occur, with the result that the net effect on accidents is small. In an attempt to evaluate this situation, a data file consisting of all the opposite-direction, head-on, and sideswipe collisions at the treatment and comparison sites was created and analyzed. It must be emphasized that the study sites were not initially chosen because of their incidence of opposite-direction crashes; in fact, if this criterion had been used in the site selection, a somewhat different set of critical sites would have been chosen. In addition, opposite-direction accidents are relatively uncommon, with the result that their frequency at the study sites is only about 20 percent that of single-vehicle ROR crashes. The before and after comparison of opposite-direction crashes for the locations painted in June 1984 is given in Table 2.

TABLE 2 OPPOSITE-DIRECTION ACCIDENTS

	FAP	FAS	Both
Treatment sites	10	9	19
Before OD accidents	43	30	73
Before accident rate	0.24	0.50	0.30
After OD accidents	16	19	35
After accident rate	0.21	0.76	0.35
Comparison sites	16	22	38
Before OD accidents	69	65	134
Before accident rate	0.18	0.32	0.23
After OD accidents	34	36	70
After accident rate	0.21	0.43	0.29

With the exception of the FAP treatment sites, the opposite-direction accident rate increased in the after period. The average increase for the treatment sites was 17 percent, whereas for the comparison sites, it was 26 percent. Using contingency table techniques, it was shown that there is no significant difference in the change of frequency of opposite-direction collisions at the treatment and comparison sites. The opposite-direction accident rate also decreased at the sites marked in July 1985, but the sample size was too small for meaningful analysis.

Achievement of "Safety" at No Cost

In general, ROR accident rates decreased at the treatment sites, but decreases of similar or greater magnitude were observed at the comparison sites. If comparison sites, which are needed to ensure reliable analyses *(22)*, had not been used, an unwary analyst might have concluded that the observed reduction in accident experience at the treatment sites was significant. The question should then be asked, What caused the reduction in accident rates at those sites that were not improved?

From 1981 to 1985, New Mexico's accident reporting threshold remained constant. The site selection process carefully eliminated sections of road that had or would experience construction activity during the study period. In response to vehicle safety standards, the overall safety of the vehicle fleet probably improved by a minor amount. There is no indication that enforcement activity or driver behavior changed between the before and after periods. During the 5-year study period, the accident experience on New Mexico's rural highways decreased slightly. Other than the painting of wide edgelines, conditions at the treatment and comparison sites were essentially unchanged, or were subject to minor changes that occur over time.

Although portions of the highway safety community are reluctant to accept the fact, it can be shown *(21)* that a set of locations selected for their high accident experience during one time period will, on the average, improve during a subsequent time period. This principle of regression-to-the-mean is not intuitive, but it can be easily demonstrated. In the initial planning for this study *(20)*, New Mexico's rural road system was divided into 7,920 1-mi segments, with termini established by milelog. The ROR accidents on these roads from 1980 to 1982 were assigned to the appropriate 1-mi segments, and the number of segments with 0, 0.33, and 0.67 accidents per year was determined. The 1983 accident experience on these sections, few of which were improved, was also determined. For example, the 94 segments that averaged 2.0 accidents per year in 1980 and 1982 had the following accident experience in 1983:

ROR accidents, 1983	0	1	2	3	4	5
Number of segments	31	36	17	3	6	1

The data show that 67 of the segments improved while 10 got worse. The average ROR accident experience on the 94 segments changed from 2.0 to 1.15, a 42 percent decrease. A similar trend was observed for all groups of roadway sections that had ROR accidents in the before period. On the other hand, the rather large group of segments with zero accidents in 1980–1982 experienced an increase; the result is expected because they could not have gotten any "safer." This same phenomenon is affecting the treatment and comparison sites selected for the wide edgeline study because of their unusually high accident experience.

CONCLUSIONS AND RECOMMENDATIONS

This study has monitored the accident experience of approximately 530 mi of rural two-lane FAP and FAS highways in New Mexico. These roadway sections were selected for study because of their unusually high rates of single-vehicle run-off-the-road accidents. In June 1984, 100 of these miles were marked with 8-in.-wide edgelines, and in July 1985, an additional 76 mi were marked. The remaining 353 mi were maintained in their normal manner and were used in this study as comparison sites. During the 5-year study period, more than 2,100 ROR accidents occurred on these sections.

The data indicate that the accident rate decreased by 28 percent at the FAP treatment sites and increased by 22 percent at the FAS treatment sites; the overall change was a –10 percent. During this same period, the comparison site rate decreased by 24 percent on the FAP and increased by 8 percent on the FAS, for an overall change of –16 percent. Clearly, the treatment sites did not perform any better than the comparison sites. At the sites marked in July 1985, the after accident rate increased at the FAP sites and decreased at the FAS sites, resulting in an overall decrease of 17 percent. The overall change in accident rate for the same time periods at the comparison sites was –24 percent. Even though the sample sizes are smaller in this latter case, the results are the same.

The single-vehicle ROR accident data were subdivided into day and night and curve and straight categories, and reanalyzed. The overall reduction in accident rates at night and on curves was similar for the treatment and comparison sites. There is no basis for concluding that the application of wide edgelines provides a benefit under conditions of darkness or curvature. The extended roadway sections treated in this project included both curves and tangents. It could be argued that the application of wide edgelines only in the vicinity of curves, while retaining standard edgelines on tangents, would be an effective spot improvement. A previous survey *(14)* found, however, that engineers believe that other treatments, including chevrons, delineators, and other warning signs, are more effective than markings for spot improvements at curves.

Finally, the incidence of opposite-direction collisions was examined. The treatment and comparison sites were not chosen initially because of the high rates of these accidents; in fact the rate of these collisions was only about 20 percent of the ROR rate. The overall opposite-direction crash rate increased at both the treatment and comparison sites. However, statistical testing showed that there was not a significant difference between the two types of sites.

The evaluation of the sites treated in July 1985 was hampered by the small sample sizes in the after period. When the 1986 accident data become available, they will be used to complete the analysis.

A previous study *(17)* suggests that wide edgelines improve the tracking behavior of motorists. However, there is no evidence from the current study that this improvement translates into a reduction in any of the accident types that this countermeasure would logically be expected to affect. Pending an evaluation of New Mexico's 1986 accident data for the second set of treatment sites, and an evaluation of the data from other states participating in the FHWA study, New Mexico will discontinue the use of wide edgelines.

ACKNOWLEDGMENT

This project was funded by the New Mexico State Highway Department; the conclusions are not necessarily those of the sponsor. Access to the accident record and roadway inventory data files used in this research was provided by the University of New Mexico Division of Government Research.

REFERENCES

1. *Fatal Accident Reporting System 1983*, NHTSA, U.S. Department of Transportation.
2. *New Mexico Traffic Accident Data 1984*. New Mexico Traffic Safety Bureau, Santa Fe.
3. J. W. Hall. Guardrail Installation and Improvement Priorities. In *Transportation Research Record 904*, TRB, National Research Council, Washington, D.C., 1983.
4. *Handbook of Highway Safety Design and Operating Practices*. FHWA, U.S. Department of Transportation, 1978.
5. *Highway Design and Operational Practices Related to Highway Safety*. American Association of State Highway and Transportation Officials, Washington, D.C., 1967.
6. *The 1984 Annual Report on Highway Safety Improvement Programs*. FHWA, U.S. Department of Transportation.
7. H. E. Ross, et al. *Crash Tests of Small Highway Sign Supports*. Report FHWA-RD-80-502. U.S. Department of Transportation, 1980.
8. H. E. Ross, et al. *Crash Tests of Rural Mailbox Installations*. Report FHWA-RD-80-504. U.S. Department of Transportation, 1980.
9. Roadside Safety Design for Small Vehicles. Proposal from the Texas A&M Research Foundation for NCHRP Project 22-6, College Station, Tex., 1984.
10. J. W. Hall. Survey of Single Vehicle Fatal Rollover Crash Sites in New Mexico. In *Transportation Research Record 819*, TRB, National Research Council, Washington, D.C., 1981, pp. 1–8.
11. P. H. Wright. Priorities for Roadside Hazard Modification. *Traffic Engineering*, Vol. 46, No. 8, Aug. 1976.
12. P. L. Zador, et al. Superelevation and Roadway Geometry: Deficiencies at Crash Sites. In *Transportation Research Record 1026*, TRB, National Research Council, Washington, D.C., 1985, pp. 43–52.
13. J. L. Graham. Effectiveness of Clear Recovery Zones. In *Transportation Research Record 923*, TRB, National Research Council, Washington, D.C., 1983, pp. 72–86.
14. P. H. Wright, et al. Low-Cost Countermeasures for Ameliorating Run-off-the-Road Crashes. In *Transportation Research Record 926*, TRB, National Research Council, Washington, D.C., 1983.
15. S. Bali, et al. *Cost-Effectiveness and Safety of Alternative Roadway Delineation Treatments for Rural Two-Lane Highways*. Report FHWA-RD-78-51. U.S. Department of Transportation, 1978.
16. *Manual on Uniform Traffic Control Devices*. FHWA, U.S. Department of Transportation, 1984.
17. Engineering the Way Through the Alcohol Haze. *ITE Journal*, Vol. 50, No. 11, Nov. 1980.
18. J. D. Brogan and J. W. Hall. Using Accident Records to Prioritize Roadside Obstacle Improvements in New Mexico. In *Transportation Research Record 1047*, TRB, National Research Council, Washington, D.C., 1985.
19. *Evaluation of Criteria for Safety Improvements on the Highway*. Roy Jorgensen & Associates, Gaithersburg, Md., 1966.
20. J. W. Hall and R. G. Ringer. Safety Effectiveness of Wide Edgelines. New Mexico State Highway and Transportation Department, Santa Fe, May 1984.
21. E. Hauer. A Common Bias in Before-and-After Accident Comparisons and Its Elimination. In *Transportation Research Record 905*, TRB, National Research Council, Washington, D.C., 1983.
22. F. M. Council, et al. *Accident Research Manual*. Report FHWA-RD-80-16. U.S. Department of Transportation, 1980.

DISCUSSION

ANITA W. WARD

Potters Industries Inc., 377 Route 17 South, Hasbrouck Heights, N.J. 07604.

Congratulations are extended to the officials at the New Mexico State Highway Department for their willing participation in an innovative approach to reduce accidents as well as the accident severity level through the use of wider-than-standard 8-in. edgelines. However, the preliminary recommendation by the author to discontinue 8-in. edgeline treatment on rural highways in New Mexico until the results of other studies are evaluated appears premature in light of the following observations.

PAVEMENT MARKINGS MUST BE VISIBLE TO BE EFFECTIVE

As Hall correctly indicates, some drawbacks of pavement markings include "deterioration due to traffic and environmental conditions, lower visibility on wet pavements, and blockage by snow and dirt" *(1)*. Hall's evaluation, however, considered neither weather nor surface conditions. Analysis of accident experience on clear road surfaces and on dry nights may prove insightful.

Pavement marking visibility at night is dependent on the retroreflective properties of glass beads embedded in the traffic paint. Although it may not have been typical of the lines applied, the 8-in. edgeline slide Hall projected to accompany his presentation indicated bead coverage only in a central 4-in. segment of the line, thus effectively providing drivers with the visual image of only a 4-in. line at night.

The State of New Mexico is making a conscious effort to upgrade the quality of its pavement markings. When the test lines were installed, however, some quality variances were noted. In 1984 one observer reported an application of 13 gal of paint per 8-in. line-mile where 32 gal should have been applied.

Another investigation noted a bead application rate of only 3½ lb/gal of paint versus a specified 6 lb/gal. Moreover, 101 mi of 8-in. edgelines were painted in June of 1984 and 76 mi were painted in July and October of 1985. Even with the highest quality of application, pavement markings have a defined service life. Because they were not restriped, it is unlikely that the pavement markings at the treatment sites remained fully effective over the life of the evaluation.

TREATMENT VERSUS CONTROL SITES

The State of New Mexico's efforts to upgrade the quality of pavement markings should have resulted in control sites that provide a stronger signal to the driver, as these controls have been repainted in the after period. Although delineation in the control sites was improving in the after period, selection criteria in the before period was more stringent for treatment sites than for control sites. As the data in Table 3 indicate, the selection process resulted in treatment sites with higher accident percentages on curves and at night.

WIDE EDGELINE SUCCESSES TO DATE

Hall's recognition that other studies of this treatment are being evaluated is a welcome indication of maintaining an open mind. A growing body of evidence indicates the potential effectiveness of wide edgelines, and it has been suggested that once factors such as environmental conditions and ADT have been taken into consideration, New Mexico may also identify candidate sites for effective treatment with wide edgelines.

In 1984, wider-than-standard edgelines were installed on more than 650 km of highway in Western Australia. The edgeline width of 150 cm was chosen as "a positive measure to reduce single vehicle accidents involving alcohol-affected or tired drivers" *(2)*. Although this is not a controlled before and after study, the author believes the results to be conservative because traffic volumes on all roads are increasing. Single-vehicle accidents were reduced by 34 percent on the treated sites, alcohol-related accidents declined by 24 percent, and the Australian Highway Department reported a benefit-to-cost ratio of 4 to 1. Other reported benefits include significant savings from a dramatic reduction in shoulder maintenance requirements and enhanced safety for cyclists *(2)*.

Ohio let contracts to paint 321.97 mi of two-lane rural state highway with wide edgelines in 1983. The lines were painted in 1983 and 1984, and Ohio has undertaken two preliminary evaluations of the after data. It is important to note that "this is a rough evaluation and is not to be quoted as a preliminary or final statement on the value of the wider edgeline" *(3)*. Based on 2 years of after data for the 8-in. edgelines installed in 1983 and 16 months of after data for the 8-in. edgelines installed in 1984, however, accident experience in the 8-in. edgeline sections either decreased further or increased less than accidents in the control sections. It is particularly significant that the greatest improvement in accident experience has been with respect to injury and fatal accidents *(3)*. Although Hall states that

TABLE 3 SINGLE VEHICLE ROR ACCIDENTS BEFORE *(1)*

	Curves			Night		
	Accidents on Curves	Total Accidents	Percentage of Total	Accidents at Night	Total Accidents	Percentage of Total
Control	503	928	54	459	928	49
Treatment	242	382	63	218	382	57

"clearly this treatment would not have an effect on the severity of the accidents that do occur" *(1)*, his contention is disputed by the actual field results to date. The effectiveness of standard width, 4-in. edgelines in reducing fatal and injury accidents to a greater degree than overall accidents has been repeatedly demonstrated in state, local, and international field tests, as well as in the Federal Highway Administration's Pavement Marking Demonstration Program evaluations *(4)*. Strengthening edgeline width from 4 to 8 in. apparently strengthens the injury mitigating potential as well.

Similar indications of the effectiveness of wide edgelines have been reported in three U.S. counties. In Los Angeles County, California, accidents were reduced by 85.7 percent when wide edgelines were installed. Because of a short test period and small sample size, the accident reduction is not considered statistically significant. However, Los Angeles County concluded that the use of 8-in. edgelines may be beneficial and recommended that additional highway sections be selected for further study of 8-in.-wide edgelines *(5)*. In Spokane County, Washington, while total accidents and injury accidents increased 12.1 and 19.3 percent, respectively, roadway sections with 8-in.-wide edgelines showed decreases of 9.2 and 32.6 percent for these accident categories. These reductions occurred even with traffic volume increases of 9.4 percent on the county road system. Accidents involving drivers who had been drinking showed even more beneficial results. County-wide, alcohol-involved total accidents and injury accidents showed increases of 7.9 and 4.8 percent, respectively. In the wide edgeline sections, alcohol-involved total accidents and injury accidents showed decreases of 28.7 and 39.4 percent, respectively *(6)*. In Morris County, New Jersey, the county engineer has adopted the practice of striping 8-in. edgelines on all county roads. Two years of annualized before and after accident data indicate that on the 115 mi of county roads where these lines were applied, dry-weather fatal and injury accidents declined by 16.1 percent, compared to a decline of only one-half that amount (8.2 percent) on other county roads in New Jersey during a comparable period. As expected, the largest percentage reduction occurred in dry weather at night, when edgelines should be most effective. After the installation of 8-in. edgelines, Morris County experienced a 21.8 percent decrease in injury accidents under dry-weather night conditions *(7)*.

In addition to the positive results indicated by these test demonstrations, at least seven U.S. jurisdictions have already adopted wide edgelines as a standard marking practice. Stronger marking patterns have been in use in Europe for years. Sound engineering judgment and accident reductions experienced to date support New Mexico's initial innovation in exploring the accident reduction potential of 8-in. edgelines, as well as Hall's recommendation to continue to further evaluate the potential benefits through other ongoing studies.

REFERENCES

1. J. W. Hall. *Evaluation of Wide Edgelines.* Report FHWA-NMSHD-86-1. Bureau of Engineering Research, The University of New Mexico, Albuquerque, N.Mex., June, 1986, pp. 3, 4, 12.
2. P. Moses. Edgelines and Single Vehicles Accidents. *Western Roads,* (Australia), April 1986, pp. 6–8.
3. Status Report on Ohio's 8-in. Edgeline Evaluation. *1986 Annual HSIP Report,* Appendix G, Ohio Department of Transportation, Columbus, pp. 4, 7, 8.
4. *1981 and 1982 Highway Safety Stewardship Reports.* FHWA, U.S. Department of Transportation, 1981, pp. 41–58; 1982, pp. 39–64.
5. J. C. Beke. Letter to California Traffic Control Devices Committee, Los Angeles, 1985.
6. R. Kelly. Preliminary Analysis of Spokane County Wide Edgelines. Washington, 1985.
7. N. D. Nedas. Letter to California Traffic Control Devices Committee, Potters Industries, Hasbrouck Heights, N.J., 1985.

AUTHOR'S CLOSURE

I would like to thank Anita Ward and Potters Industries Inc., for commenting on this paper. The discussion, however, raises a few points that must be clarified.

Because the discussion gently chides the New Mexico State Highway Department marking crews for the insufficient use of beads and paint, the reader may assume that using more paint and beads would have altered the results. I have been advised by the highway department that the 8-in. markings were applied in the same manner as 4-in.-wide lines, except that two paint guns were used. To be precise then, the study essentially compares the typical 4-in. line installed by the New Mexico State Highway Department with a line that is twice as wide, using paint and beads at twice their standard application rate.

In my paper I have outlined in great detail the procedure used to select candidate sites. In summary, 150 sections of rural highway with unusually high ROR accident experience, as determined by the critical rate technique, were identified. This set of candidate sites was subsequently pared by eliminating short sections as well as those on multilane highways and those scheduled for improvement. In order to avoid the detrimental effect of small sample sizes, treatment sites were not chosen because of their accident experience on curves or at night, or for the amount of travel under these conditions. It is not surprising, therefore, that the treatment and control sites exhibit differences in their distributions of accidents on curves and during the hours of darkness. It should also be noted that the comparison sites averaged greater lengths and lower volumes than the treatment sites. Although having identical values for these characteristics at the two types of sites may be reassuring, it is certainly not essential.

Proponents of edgelines in general and wide edgelines in particular contend that these devices have a greater effect at night and on curves. The analysis reported in this paper was unable to support this contention. It appears contradictory, however, for Ward to protest that the treatment sites had a higher percentage of accidents under these conditions, which hypothetically provide a greater opportunity for improvement.

The discussion also faults the research for failing to evaluate the effect of wide edgelines on dry pavement during the hours of darkness. As previously noted, because neither darkness nor weather were considered in the site selection process, it is quite possible that a different set of treatment and control sites would have been chosen if these criteria had been used. Although the comparison sites should provide an adequate control for darkness, there is no guarantee they properly control for variations in weather. For these reasons, these characteristics were not discussed in the paper. In the interest of completeness, I will report that in the before period, 42 percent of the treatment site accidents and 38 percent of the comparison site accidents occurred on dry pavement at night. The comparable figures in the after period were 34 and 33 percent, respectively. It is not possible to assign any practical or statistical significance to these values.

Numerous studies reported in the technical literature document the ability of selected countermeasures to reduce crash severity. For example, it is generally agreed that medians, impact attenuators, and obstacle removal can reduce the severity of ROR accidents. The mechanisms that contribute to ROR accident severity reduction are easily identified; however, it would be difficult to argue that they reduce the incidence of roadside encroachment. Conversely, if wide edgelines provide the guidance suggested by their proponents, they could reduce the incidence of roadside encroachments, but there is no logical basis for concluding that they disproportionately affect injury accidents. The severity of ROR accidents is principally determined by speed, the nature of the roadside, vehicle type, restraint usage, and similar factors. It is not influenced by the width of the edgeline, and I seriously question any study that concludes otherwise.

Although Ward references a number of studies that have concluded that there may be a benefit associated with the application of wide edgelines, several of these studies are, in fact, fatally flawed by their failure to use control sites. Regression-to-the-mean is discussed; to reiterate, sites chosen for their perceived hazard in one time period may subsequently "improve" even in the absence of any treatment. Failure to monitor the changes over time on a set of hazardous control sites will result in an overestimation of treatment benefits. The use of large sample sizes, the adjustment for changes in traffic volume, or the use of systemwide accident experience as a control, are clearly not sufficient to alleviate this problem. The reader may want to compare the 16 percent reduction in control site accident experience found in this study with the results cited by Ward.

Although I recognize that the findings of this limited study differ from those of earlier research and, to some extent, may contradict "conventional wisdom," I did not intend to attack either the previous studies or the researchers who have conducted them. In contrast, my purpose was purely a cautionary one urging traffic engineers who may be considering this treatment to proceed circumspectly. If wide edgelines have positive effects, they can only be detected through a carefully controlled study with a large number of sites that are maintained and monitored over an extended time period.

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

Weight-Specific Highway Sign Effects on Heavy Trucks

FRED R. HANSCOM

The objective of this study was to test the field effectiveness of the Grade Severity Rating System (GSRS), via application of weight-specific signs to control truck speeds on downgrades. Before-after sign effects were evaluated in terms of speed differences and incidences of smoking brakes for trucks in specific weight categories. A five-state (California, Colorado, Idaho, Oregon, and West Virginia) sample of study sites included grades of varying severity. The study design included a determination of novelty effect as well as concurrent before-observations at selected control sites (i.e., no weight-specific sign present). In addition, the feasibility of state highway agencies' conducting an accident study to assess GSRS safety impact was examined. Weight-specific signing (WSS) was determined to elicit a favorable before-after effect at high-severity sites. Three truck behavioral measures provided the basis for this result: mean truck speed, percentage of trucks exceeding posted WSS speeds, and incidences of smoking brakes. Before-after reductions in mean speed were observed at two out of three high-grade-severity locations following installation of the WSS. Substantiating evidence that the WSS was responsible for the speed reduction evolved from (a) corresponding speed increases at one matched control site and (b) the absence of speed changes for trucks weighing less than 70,000 lb (31.8 Mg) at the other site. Percentages of trucks exceeding WSS-posted speeds were reduced for 70,000 to 80,000 lb (31.8 to 36.3 Mg) trucks at one site and for 60,000 to 70,000 lb (27.2 to 31.8 Mg) trucks at the other. The proportion of trucks characterized by smoking brakes was reduced at the single high-severity site where this measure was observed. Because GSRS represents the state of the art, its application was viewed to improve states' liability positions. Weight-specific signing was recommended for use at high-grade-severity locations.

The Grade Severity Rating System (GSRS) is a technique used for reducing the incidence and severity of truck downgrade accidents. GSRS feasibility has been examined via the development and prototype application of a weight-based truck speed selection model *(1, 2)* in recent work conducted for the Federal Highway Administration (FHWA). The model was based on an empirical determination of brake heating characteristics as a function of gross truck weight, grade length, and steepness. Field application of the GSRS involves use of weight-specific signs (WSS) advising truckers of the appropriate descent speed according to gross truck weight. Figure 1 shows the GSRS by (a) defining grade severity ratings (GSR 1 through 10) and (b) prescribing safe downgrade speeds for 80,000-lb (36.3 Mg) combinations according to grade geometry.

Transportation Research Corporation, 2710 Ridge Rd., Haymarket, Va. 22069.

OBJECTIVE

The objective of this field study was to evaluate the field effectiveness of the GSRS, via application of weight-specific signs, to control truck speeds on downgrades *(3)*. A five-state sample of study sites included grades of varying severity. Before-after sign effects were evaluated for trucks in specific weight categories in terms of speed differences and incidences of smoking brakes. The study design included a determination of novelty effect as well as concurrent observations at selected control sites, that is, no weight-specific sign present. In addition, the study examined the feasibility of state highway agencies' conducting an accident study to assess GSRS safety impact.

STUDY PROCEDURE

Designation of Measures of Effectiveness

Measures of effectiveness (MOEs) refer to that which is measured in an evaluative study. Designation of MOEs derived from the primary intent of the current study: a traffic operational evaluation of WSS sign characteristics as determined by application of the GSRS. In addition, this study examined the feasibility of an accident-based evaluation.

Two operational MOEs possess high face value because of the nature of brake-fade truck accidents: (a) smoking brakes and (b) speed characteristics. Smoking brake occurrences were assessed as a proportion of total truck volume. Speed characteristics were addressed by truck weight class targeted on the weight-specific signs. Within each weight category, the before-after sign impact was determined for both the mean speeds and the proportion of trucks exceeding the posted weight-specific speed. An obviously favorable safety implication would result from reduced overall speeds and fewer violations in the "after" condition.

Study Design

Based on available site characteristics (e.g., required downgrade steepness, available truck weight data), the current study used a before-after with control site paradigm to the extent possible. Sites were designated in order to support multiregional data within the United States. Although the majority of required geometric conditions were located in the Western United States, data were also gathered at one east coast site. In order to render a precise geographic effect response to the

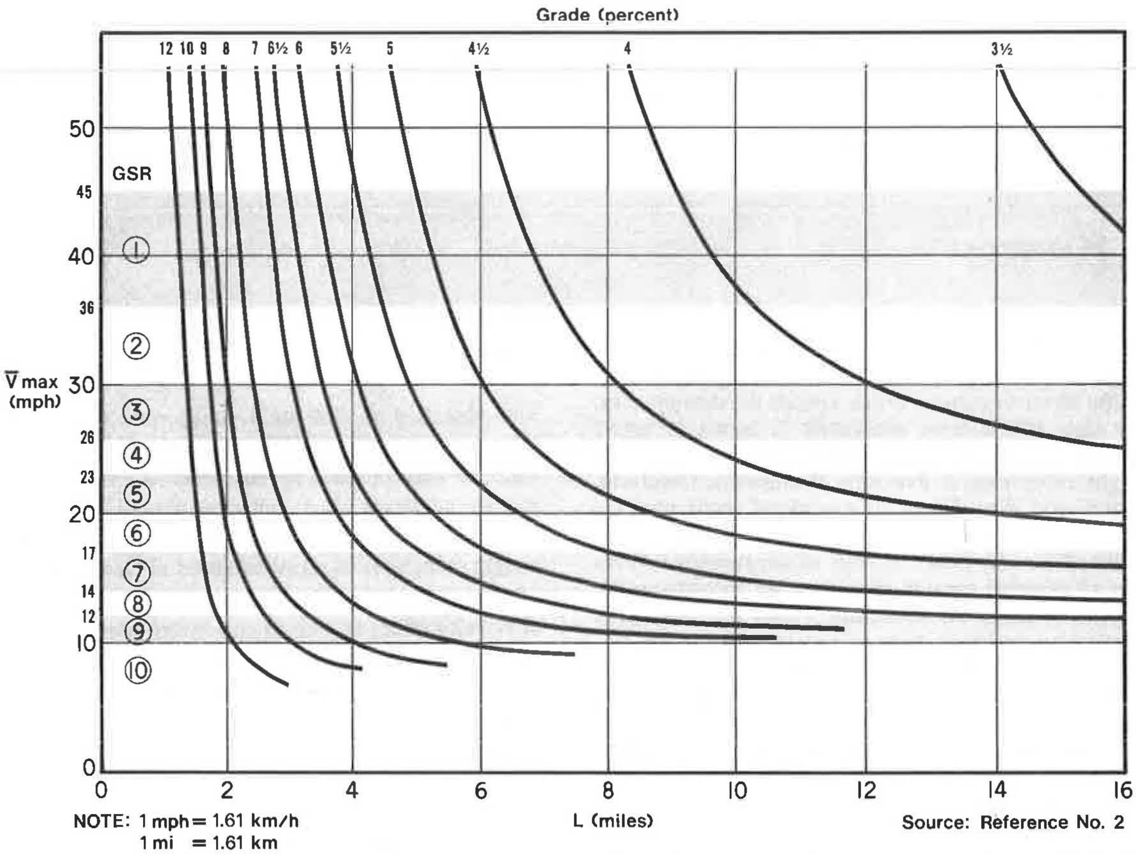


FIGURE 1 Description of GSRs in terms of grade severity ratings (GSR number) and prescribed speeds for 80,000-lb (36.3 Mg) trucks according to grade characteristics.

GSRS, the site paradigm included a closely matched geometric site pair comprised of the eastern site and a western site. Acclimation data were gathered immediately following sign installation at three sites in various areas of the United States to observe any novelty effect associated with WSS sign responses.

Site and Sign Characteristics

To achieve the required multiregional effect, the following sites were represented in the data base:

- I-5: Siskyou County, California.
- I-70: Georgetown, Colorado.
- US 95: Lewiston, Idaho.
- I-84: Cabbage Hill, Oregon.
- I-5: Medford, Oregon.
- US 40: Morgantown, West Virginia.

Both low- and high-severity sites were included in the sample. Site geometrics (e.g., no upgrades nearby) were such that observed speeds were not confounded by factors other than the downgrade. Figure 2 shows studied weight-specific signing characteristics, site designations, and highway geometric conditions that characterized each site.

Field Data Collection

Two field procedures were conducted. Manually timed speed data were collected at a point on each grade where any brake-fade speed effect (e.g., runaway truck) could be observed. In addition, truck weight data were recorded at a nearby weigh station. Each collection procedure involved recording truck-specific descriptive data used to associate individual speeds and weights. The following techniques were applied for each of three data types: truck descriptions, weights, and speeds.

Truck descriptions. A procedure was developed by which field observers could quickly extract sufficiently detailed visual truck characteristics in order to identify target trucks. Carrier name (and unit identification number, in cases in which multiple trucks from a given line were traveling in proximity) was the most helpful information in the matching procedure, which proved to be quite effective; approximately 95 percent of measured speeds and weights were matched.

Weights. Observers stationed at state-operated weigh scales recorded truck descriptions and weight information gathered by state personnel. This source of weight data provided a high level of accuracy.

Speeds. Observers worked in teams; the primary responsibility of one team member was to manually time speeds and the responsibility of another was to record truck descriptive

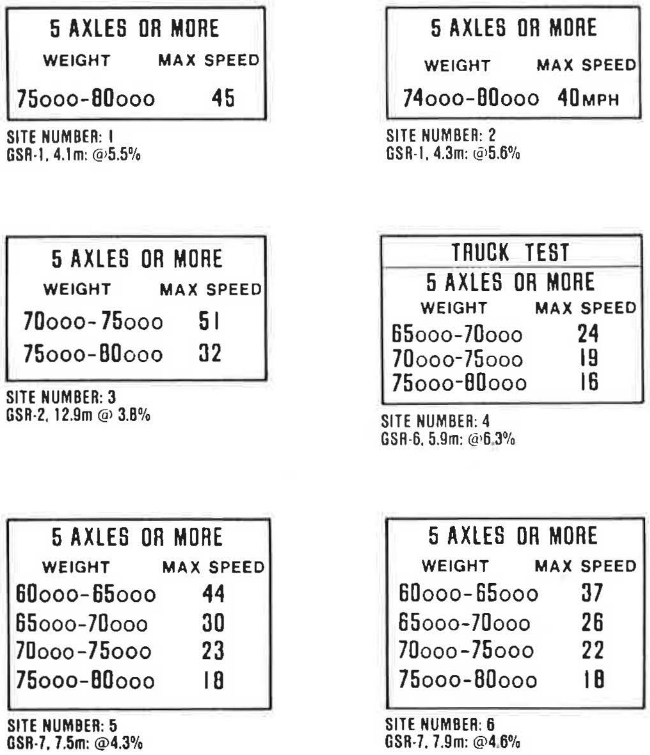


FIGURE 2 Weight-specific sign characteristics, site designations, GSR ratings, and approximate geometrics applied at six study sites.

data. This procedure was effective in producing a high capture rate. That is, when clusters of trucks appeared at the speed site, each observer was capable of gathering both speed and descriptive data; therefore, data attrition was minimized.

Manual timing was designated as the speed collection technique because it was less obtrusive than radar, which is frequently detected by truckers. The applied speed measurement technique involved manually timing target trucks between pavement markings spaced at 268 ft (81.7 m). Measurement accuracy was enhanced by the use of digital stopwatches that were capable of displaying measured time to 1/100 sec. Intercoder reliability determinations verified sample speed measurement accuracy of 0.5 mph.

RESULTS

Speed Effects

The applied field study paradigm supported the following analysis conditions:

- Before-after comparisons for all tested signs, and
- Acclimation (e.g., novelty effect) study at four sites.

Findings for each of these analysis conditions are discussed next.

Before-After Effects

A summary of observed effects for each grade severity rating (GSR) category at all six test sites is given in Table 1. Each cell in the table contains after-minus-before values for both mean speed and violation percentage; that is, trucks exceeding the posted speed limit. Statistical significance, noted by arrows depicting directionality (e.g., arrows indicate observed speeds statistically exceed WSS specified), is based on application of the Student's *t*-test for mean differences and the *z*-test for proportions. Sufficient samples supported use of the 0.01 level of significance to be applied in most cases.

A before-after data comparison (without regard to statistical significant test results) notes a predominant reduction in both mean speeds and violation percentages for heavy trucks in the "after" condition. The single exception is Site 1, the least severe grade. Specific observations of degraded operational performance at the remaining sites are as follows: Increased violation percentage at Site 2, higher mean speed at Sites 3 and 4, and increases in Site 6 data cells. Only one of these conflicting findings (violation percentage at Site 2) may be attributed to a small sample (N = 10).

Statistical significance is noted by arrows within the cells. Arrows indicate changes in mean speeds between before-and-after conditions. The predominant statistical effect is significance associated with lighter weight truck groupings (not affected by the WSS), which comprise a major portion of the sample. With the exception of Site 4, these lighter trucks exhibit lower mean speeds in the after condition. The sustained speed reduction across all weight categories noted at Site 5 is

TABLE 1 OBSERVED BEFORE VERSUS AFTER DIFFERENCES IN MEAN SPEEDS AND PROPORTIONS OF TRUCKS EXCEEDING WSS SPEEDS

	Site Designation (severity)											
	1 GSR-1		2 GSR-1		3 GSR-4		4 GSR-6		5 GSR-7		6 GSR-7	
	mph	%	mph	%	mph	%	mph	%	mph	%	mph	%
Non-GSR speeds	+1.6↑	+9↑	−2.1↓	−6↓	−1.4↓	−9↓	−0.6	0	−3.7↓	−14↓	−3.9↓	−9↓
WSS-affected speeds[a]	+4.3↑	+15↑	−5.0	+30	+2.3	−5	+2.7	+4	−3.3↓	−18↓	−1.5	−6
					−1.8	+11	−5.3↓	−16↓	−9.7↓	−10↓	+1.0	+4
							−3.3↓	−3↓	−6.0↓	−2	−0.2	+4
									−4.5↓	−1	−1.0	−1

NOTES: Arrow (↑) indicates statistically significant. Metric equivalence: 1 lb = 0.454 kg, 1 mph = 1.62 km/h.

[a]Speeds differ between sites; see Figure 2.

associated with the largest sample obtained at any of the sites.

Assessment of WSS effectiveness based on significance testing of the data is as follows. The WSS apparently alerted truckers to downhill brake-fade accident potential, as evidenced by reduced speeds for the lighter trucks as well. However, considering the target heavy-truck population, responses to the WSS were not uniform across all sites. Statistically reduced speeds for GSR-affected trucks were observed at only two of the six test sites. Despite the impressive speed reduction response to the Site 5 sign (control data gathered at a matched site confirm the sign's effectiveness), the Site 1 sign (exposed to virtually the same trucker population) was not shown to be effective. A plausible explanation for this difference is the significantly higher grade severity at Site 5.

The data indicate that the WSS is not effective at lower severity sites. No (statistically significant) speed-reducing effect was observed for heavy trucks at Sites 3 and 4, and an actual speed increase occurred at Site 1.

With regard to the high-severity GSR sites, significant truck speed reductions were observed at two of the three sites. Although trucks in all weight classes slowed at Site 5, only those heavier than 70,000 lb (31.8 Mg) were affected at Site 4. Although observed changes in speed parameters were modest [mean speed reduction ranging from 1.0 to 9.7 mph (1.6 to 15.6 km/h) and decreased speed violations ranging from 1 to 16 percent], associated statistical significance is interpreted as evidence of WSS effectiveness. Further, a matched control site in the vicinity of Site 5 experienced speed increases that further substantiates the interpretation in this instance. Although no matched control site was available in the vicinity of Site 4, it is noteworthy that unaffected trucks [i.e., those with gross weights less than 65,000 lb (29.5 Mg)] did not slow on the grade. This implies that no extraneous explanation existed to cause slowing.

It is difficult to explain between-site difference, which could account for the speed-reducing WSS impact at Sites 4 and 5 and yet result in no effect at Site 6. Sign installations were similar across all three high-severity sites and were constructed in conformity with the FHWA-specified design (see Figure 3). Factors that logically refute any expected sign-related speed-effect difference between sites are as follows. Two signs were installed (one at the top and one part-way down the grade) at all three sites. Although certain preparatory signing (e.g., a series of large yellow signs give advance warning of the downgrade) may have competed for driver attention at Site 6, large yellow grade-advance warning signs were also present at Site 4. Similarly, because of the geographic proximity of these two signs, no regional effect was found to exist in driver response.

Nevertheless, three factors unique to the Site 6 grade were noted, which may have accounted for a reduced WSS speed-reducing effect noted by the principal investigator. First, more advance grade warning signs existed at Site 6 than at any other test site. These signs (e.g., typically "first warning, steep downgrade ahead") may have diverted driver attention from the initial WSS. Second, the later WSS was slightly laterally displaced from the roadway (because of a fill slope), and driver observation of the sign may have been slightly impaired; however, the initial WSS was highly conspicuous. Finally, a number of logging trucks (operated by a variety of local companies) were noted, which consistently descended the grade at high speeds. These trucks appeared to be in good mechanical condition, and the drivers were obviously quite familiar with the roadway.

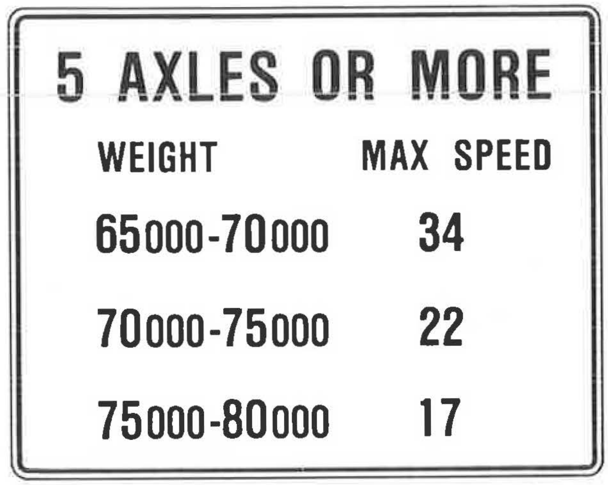


FIGURE 3 FHWA-specified design.

Acclimation Effect

To determine whether weight-specific signs elicited a novelty effect, data were gathered immediately following sign installation at three locations. Acclimation data were gathered at Sites 2, 3, and 6. Observed acclimation speed effects (Table 2) are briefly discussed for each site.

Site 2. Although certain speed differences were observed between the before and acclimation periods, those differences could not logically be attributed to appropriate sign responses. Trucks weighing less than 74,000 lb (33.6 Mg) (not addressed in the sign message) exhibited a significant reduction in mean speed and reduced proportion exceeding the posted speed; however, this effect was offset by speed increases exhibited by trucks weighing more than 74,000 lb (33.6 Mg). Speed increases for the heavier trucks were not significant because of an inadequate sample. Nevertheless, an interpretation of these data indicates no favorable acclimation effect of the weight-specific signing.

Site 3. Very slight speed differences were noted between the before and acclimation conditions. A single statistically significant effect was an increased proportion (64 versus 57 percent) of trucks weighing less than 70,000 lb (31.8 Mg) (thus not affected by the WSS) that exceeded 55 mph during the acclimation period. Therefore, no WSS-related acclimation speed effect was evident.

Site 6. Trucks not affected by the WSS [i.e., those lighter than 60,000 lb (27.2 Mg)] exhibited lower mean speeds, and a smaller proportion exceeded the 55 mph limit immediately following installation of the WSS. Although trucks in the intermediate weight classes [e.g., 60,000 to 75,000 lb (27.2 to 34.0 Mg)] demonstrated a tendency toward lower speeds, the effect was not statistically significant. Particularly noteworthy is the heaviest truck category [75,000 to 80,000 lb (34.0 to 36.3 Mg)] in which nearly the same proportion (91 and 90 percent)

TABLE 2 BEFORE VERSUS ACCLIMATION DIFFERENCES IN MEAN SPEEDS AND SAMPLE PERCENTAGES EXCEEDING POSTED SPEEDS

GSR Weight Category	West Virginia		Colorado		Cabbage Hill		Imperial Grade	
	mph	%	mph	%	mph	%	mph	%
Non-WSS affected	−3.0↓	−15↓	+1.1	+7↑	−4.4↓	−11↓	−3.4	+7
	+7.9	+30↑	−0.2	−13	−7.0	−24		
			−1.4	0	−3.1	+5	−5.0	+8
					−3.7	−4	−4.2	−37
					−1.8	−1		

NOTES: Arrow (↓) indicates statistical significance. Metric equivalence: 1 lb = 0.454 kg, 1 mph = 1.62 km/h.

exceeded the GSR-posted speed between the before and acclimation periods. Therefore, no WSS-related acclimation speed effect was evident.

In summary, consistent observations at three grades (low-, intermediate-, and high-severity sites) indicate that the WSS elicited no novel speed-reducing effect. Those before-after speed differences noted previously (with control for season of the year) were apparent effects resulting from a learned response because the signs had been in place for a sufficient duration.

Smoking Brake Effects

A secondary measure of WSS effectiveness was the incidence of smoking brakes. This behavior was so designated because, as truck brakes heat up, detectable odor and smoking comprise a warning of actual brake loss. Sufficient data samples were obtained at one intermediate-severity site (Site 3) and one high-severity site (Site 4).

Site 3

Extensive observation of Jake brake usage (i.e., utilizing engine compression to reduce speed) and incidences of smoking brakes were conducted at Site 3. Comparisons of before-after results were based on a sample of 1,476 trucks over an observation period of 9 days. The summary result is as follows:

	Before *N = 960* *(%)*	*After* *N = 516* *(%)*
Jake brake usage	30.5	33.7
Smoking brakes	11.8	15.1

Slight but statistically nonsignificant increases were noted for both Jake brake usage and smoking brake occurrence. An explanation of this effect was sought on the basis of possible differences in sampled weight distributions between the before and after conditions. Although a slight increase in heavier trucks (25 percent versus 22 percent targeted by the WSS) characterized the after study sample, this difference alone was insufficient to account for the increase in observed brake effects.

In order to assess WSS effectiveness on the basis of these measures, a slight increase in Jake brake usage and a significant reduction in smoking brake occurrences can be expected. In this case, the obviously more significant measure is smoking brake occurrences. The observed increase in the percentage of smoking brakes indicates a poor response to the WSS. This finding is consistent with the speed effect noted earlier asserting that the Site 3 WSS installation was not effective.

Site 4

Because of specialized personnel requirements to assess Jake brake usage, this measure could not be obtained at Site 4. However, observations of smoking brake incidences indicated a significant reduction as follows:

	Before *N = 595* *(%)*	*After* *N = 590* *(%)*
Smoking brakes	3.5	1.4

A check on before-versus-after weight distribution (i.e., heaviest GSRS category; 47 percent before, and 39 percent after) would account for a minimal reduction in smoking brake incidences in the after condition. Therefore, the observed before-after reduction in the proportion of smoking brake incidences is an indication of WSS effectiveness.

That smoking brake differences revealed an effect at Site 4 but not at Site 3 is consistent with observed speed effects. These findings, based on separate measures, confirm WSS effectiveness on high-severity grade.

SUMMARY AND CONCLUSIONS

Weight-specific signing was determined to elicit a favorable before-after effect at most high-severity sites tested. Three truck behavioral measures that provided the basis for this result were mean truck speed, percentage of trucks exceeding posted WSS speeds, and incidences of smoking brakes.

Modest reductions in before-versus-after mean speed were observed at two out of three high-severity locations following installation of the WSS. Substantiating evidence that the WSS was responsible for the speed reduction evolved from (a) corresponding speed increases at a matched control site, and (b) the absence of speed changes for trucks weighing less than 70,000 lb (31.8 Mg) at the other site. Percentages of trucks exceeding WSS-posted speeds were reduced for 70,000- to 80,000-lb (31.8- to 36.3-Mg) trucks at one site and 60,000- to 70,000-lb (27.2- to 31.8-Mg) trucks at the other. At the third high-severity site, higher speeds were observed for a subsample of truckers who were quite familiar with the grade. The proportion of trucks characterized by smoking brakes was reduced by one-half at the single high-severity site where this measure was observed.

A final consideration is the issue of states' liability. Although detailed study of liability implications of WSS installations was beyond the scope of the existing contract, it is nevertheless a concern. Litigation against states may occur in the event of brake-fade accidents. Two related viewpoints held by states were brought to the author's attention during the course of this study. The first is that weight-specific signing is superior to conventional advisory truck speed limits on downgrades in that, because of its greater specificity and conspicuity, it is likely to result in a safety benefit. Therefore, a state's legal position would be improved as a result of WSS application. The second is that, assuming compliance with WSS-posted speeds, greater stream flow perturbation would result from speed differentials between trucks of varying weight, and safety would be degraded. Therefore, a state's legal position would be weakened as a result of WSS application.

Consideration of WSS liability implications is as follows. Although before-and-after speed reductions were frequently observed following WSS application, the overall slowing effect was not of sufficient magnitude to increase intervehicle speed differentials. More important, the liability issue could best be resolved by assessing whether the state acted prudently when signing the downgrade. Because the GSRS comprises the state of the art in reduction of brake-fade accidents, and has in this study proven to be somewhat operationally effective, the conclusion is that states' liability position would be improved by the use of WSS.

Although actual significant speed reductions were observed at only two of six test sites, this finding is considered a basis for recommending WSS application for the following reasons. First, the signs demonstrated greater effectiveness in the presence of the more severe hazard, a finding that substantiates both sign credibility and safety effects. Second, although actual observed mean speed reductions were slight (3.3 to 9.7 mph), their statistical significance attests to their efficacy at driver behavior modification. Finally, as noted previously, WSS usage provides a liability-protection benefit to state highway agencies.

RECOMMENDATIONS

Reductions in truck speeds and smoking brake occurrences at certain high-severity grade sites were observed in this study. Further, it was concluded that GSRS application would improve a state's liability position because the GSRS comprises the state of the art in brake fade accident prevention.

Therefore, application of weight-specific signing is recommended at high-severity grade locations (i.e., GSR 6 or above). Specific geometric conditions comprising a GSR 6 grade are as follows:

Percent	*mi*	*km*
4.5	14.0	22.5
5	12.0	19.3
5.5	6.6	10.6
6	5.2	8.4
6.5	4.4	7.1
7	3.8	6.1
8	3.0	4.8
9	2.4	3.9
10	2.0	3.2

Further research to improve driver compliance with weight-specific speeds is also recommended. Application of automatic weight sensors in pavements that are integrated with changeable message bulb-matrix signing offers the potential for increased compliance by providing highly conspicuous speed information on a truck-specific basis.

ACKNOWLEDGMENT

The work reported in this paper was sponsored by the Federal Highway Administration, U.S. Department of Transportation and monitored by George B. Pilkington, II.

REFERENCES

1. T. M. Myers et al. *Feasibility of a Grade Severity Rating System.* Report FHWA-RD-79-116. FHWA, U.S. Department of Transportation, Aug. 1980.
2. W. A. Johnson et al. *The Development and Evaluation of a Prototype Grade Severity Rating System.* Report FHWA-RD-81-185. FHWA, U.S. Department of Transportation, March 1982.
3. F. R. Hanscom. *Field Test of the Grade Severity Rating System.* Report FHWA-RD-86-011. FHWA, U.S. Department of Transportation, Sept. 1985.

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

Regional Differences in Preferences for Median Crossover Signing

GILLIAN M. WORSEY, CHARLES E. DARE, AND RICHARD N. SCHWAB

Described in this paper is a study of advance warning signs for median crossovers on divided highways. Candidate crossover signs were identified from a literature review, survey of current state practices, and discussions with FHWA personnel. Seven of these signs were selected for further testing in a laboratory study for legibility, understanding, and driver preference. Sixty subjects representing a cross-section of drivers participated in the study: 30 at the Turner-Fairbanks Highway Research Center in McLean, Virginia, and 30 at the University of Missouri-Rolla in Rolla, Missouri. Two of the seven signs were word messages and five were symbolic signs. The results from both groups of participants showed that the most appropriate word message sign would appear to be "median crossover." This sign was understood best by the participants to whom it was shown, and "crossover" was the word the majority of participants believed best conveyed the intended meaning. Of the symbolic signs tested, the one found to be the best was that of two median noses. This symbolic sign performed well in tests of legibility and understanding and was the sign least often confused with other signs. It was also the symbolic sign most preferred by the participants and was the simplest of the symbolic designs. The symbolic signs were substantially more legible than the word messages, and the symbolic design of two median noses is recommended to identify median crossovers.

Median crossovers are often provided on divided highways between intersections for the use of emergency vehicles and to accommodate minor turning movements for convenient access to adjacent roadside development. About 35 percent of the accidents that occur between intersections on four-lane highways involve median openings *(1)*. As a result concern has been expressed that public-use crossovers may be hazardous, especially where visibility of the crossover is limited. If used, such crossovers should be signed to provide advance warning to drivers. Hazards associated with crossovers include (a) vehicles slowing down in the fast lane of a divided highway or accelerating into it, (b) vehicles turning across the divided highway, and (c) vehicles making sudden lane changes. These maneuvers may possibly lead to rear-end or broadside collisions.

The third revision of the *1978 Manual on Uniform Traffic Control Devices* (MUTCD) *(2)* provides for the use of a median crossover sign (D13-1, see Figure 1) but this is a large (6- × 3-ft) guide sign and there may not be sufficient room on suburban divided highways to erect such a large sign. Also, it is not the color that is customarily used for warning messages. The MUTCD does not currently suggest an advance warning sign for median crossovers, although it does suggest that a green and white advance message sign showing the distance to the crossover (D13-2) may be used.

The principal findings of a study to determine the most appropriate design of an advance warning median crossover sign are discussed in this paper. The objective of the study was to identify alternative designs for median crossover signs from a nationwide review of practices for signing median crossovers and related literature on traffic signs. These alternative designs were then tested for legibility, recognition, meaning, and preference. They were first tested at the Turner-Fairbanks Highway Research Center in McLean, Virginia, and later at the University of Missouri-Rolla, thus enabling a comparison to be made between the results obtained in Virginia and those obtained in Missouri.

METHOD

Participants

The Virginia participants were paid volunteers recruited from among research fellowship students and computer center staff at the Turner-Fairbanks Highway Research Center and from a list of participants in previous experiments at the center.

Thirty participants were tested, 10 (5 males and 5 females) in each of the following age groups: 17 to 29, 30 to 49, and 50 and over. The mean age of participants in each group was 22.6, 40.4, and 58.6 years, respectively. All participants had their vision tested on an Ortho-Rater to ensure corrected visual acuity of 20/33 or better and to ensure normal color vision. The mean visual acuity was 20/20.

The Missouri participants were unpaid volunteers recruited from among psychology and civil engineering students, staff, faculty, and wives of faculty members at the University of Missouri-Rolla. Thirty subjects in the 17 to 29, 30 to 49, and 50

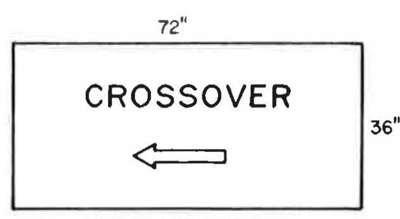


FIGURE 1 Median crossover sign (D13-1) ***(2)*****.**

G. M. Worsey and C. E. Dare, Department of Civil Engineering, University of Missouri-Rolla, Rolla, Mo. 65401. R. N. Schwab, Federal Highway Administration, Turner-Fairbank Highway Research Center, 6300 Georgetown Pike, McLean, Va. 22101.

and over age groups were tested. The mean age of participants in each group was 20.7, 41.2, and 58.3 years, respectively. The differences in mean ages for the Missouri and Virginia participants were fairly small.

The only method available for testing to ensure corrected visual acuity of 20/30 or better was a Snellen Eye Chart, which only allowed visual acuity to be classified as 20/20 or 20/30. Unfortunately, color vision could not be tested but their color vision was correct according to each participant and no one had problems with colors during the experiment.

Apparatus

Seven candidate signs for median crossovers were studied in the experiment. These included five symbolic designs and two word signs. The design of the signs came from several sources, including a survey of state highway departments (two signs), a literature review (one sign), FHWA personnel (two signs), and a Virginia crossover sign. The word signs included "crossover," as this is the wording on the signs in Revision 3 of the *1978 Manual on Uniform Traffic Control Devices*, and "median opening." Questions about wording were included in the last part of the experiment. The 7 signs along with the 13 distractor signs used in the experiment are shown in Figure 2. Nine other sign designs (from the same sources) were considered, but in order to keep testing time to approximately 1 hr, only a limited number of signs could be tested. The other designs considered are shown in Figure 3. When time was available, the "median crossover" sign was shown to the Missouri participants.

At the suggestion of the FHWA Office of Traffic Operations, all the signs tested were black on yellow diamond warning signs, with the exception of the Virginia crossover sign and the permissive U-turn sign suggested by the Office of Traffic Operations. Instead of a green ring to denote a permissive sign as has been tested in previous sign studies *(3)* the Office of Traffic Operations suggested using a green periphery (see Figure 2). The signs were composed on a computer graphics system and superimposed onto a digitized photograph of a median crossover from which slides were made.

Thirteen signs were used as distractors. These included a permissive right-turn sign, similar to the permissive U-turn sign, and a railroad crossbuck outlined in red, which was part of another FHWA study. Of 11 signs from the MUTCD, 10 were chosen because they had already been drawn on the computer graphics system. A type 3L object marker was also used because the Virginia crossover sign was similar in size to

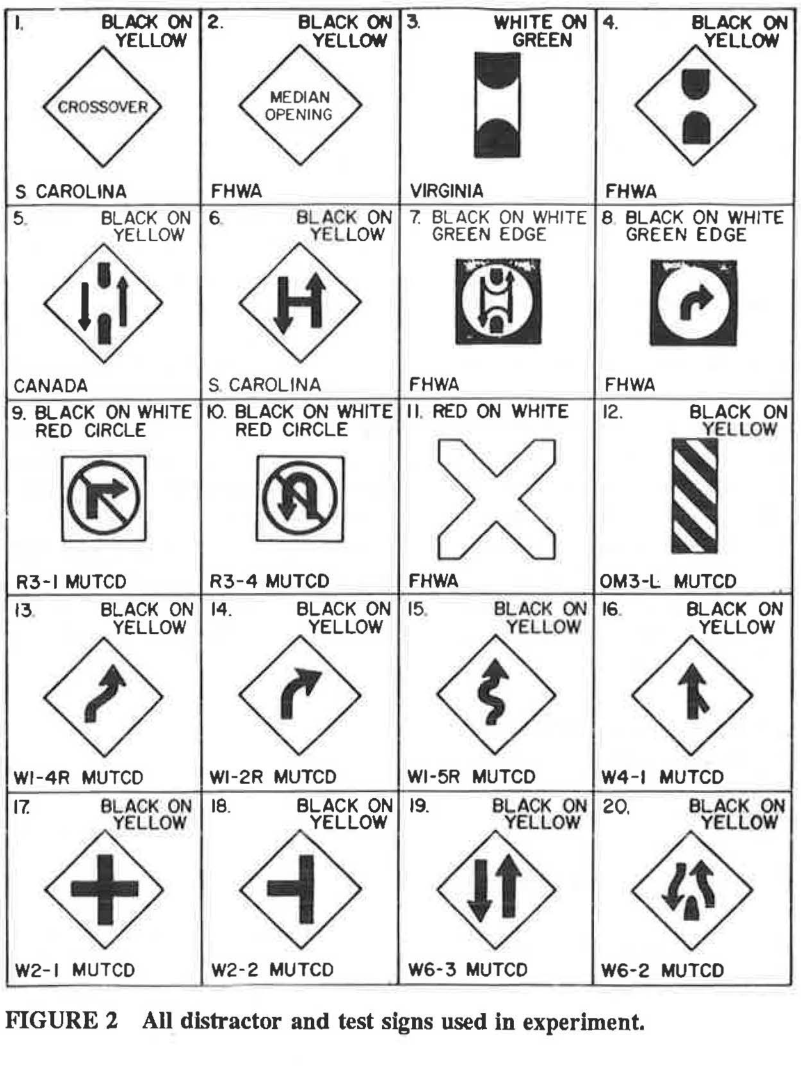


FIGURE 2 All distractor and test signs used in experiment.

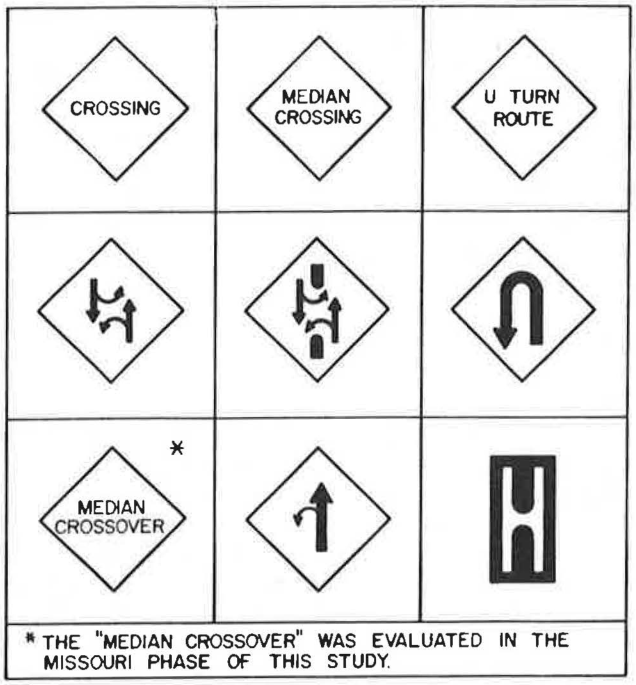


FIGURE 3 Additional signs considered for use in experiment.

an object marker. All distractor and test signs used are shown in Figure 2.

The slides were rear-projected onto a translucent screen. The size of the projected image of the signs was 2-3/8 in. from point to point of the yellow diamond in a 16-5/8- × 10-1/2-in. background scene of a median crossover. This size was chosen so that the participants with the best eyesight, although they could see a yellow and black sign, for example, could not recognize the meaning of familiar signs at the farthest distance from the image (110 ft). This was purely a laboratory experiment and no attempt was made to relate the distances measured to equivalent distances for standard-sized signs.

The experiment was conducted in a concrete tunnel approximately 12 × 12 × 120 ft underneath the structures laboratory at the Turner-Fairbanks Highway Research Center in McLean, Virginia. The slide projector and hanging screen were set up at one end of the tunnel.

The same set of slides was shown to the Missouri participants with the addition of "a median crossover" sign. The slides were again rear-projected using exactly the same type of slide projector as that used in Virginia except that it had a smaller screen and stood on a table. A facility equivalent to the tunnel in Virginia was not available so the test was conducted in the third floor corridor of the civil engineering building at the University of Missouri-Rolla.

Procedure

The participants completed a biographical data and consent form first. If they wore corrective lenses for driving, they also wore them during the test. The same procedure was followed in Missouri and Virginia.

Legibility and Meaning

The instructions for Parts I and II of the test were read to the subjects. After answering any questions the participants might have had, the examiners presented the first slide on the screen. The participants walked individually toward the projected sign until they could identify any feature on the sign. The feature and the distance at which it was identified were recorded. This procedure was repeated until all the major features of each sign had been identified.

The participants were also instructed to give the meaning of the sign as soon as they believed they knew what it meant. If they gave the wrong meaning, they were instructed to try again, and their error was recorded as a misinterpretation.

When all of the features of the sign had been identified, the participant walked back to the 110-ft mark, the next slide was presented, and the procedure was repeated. This process was repeated until the participant had seen all 20 slides. The slides were presented in random order (which was different for each participant) with the proviso that the first two signs were not crossover signs. In this way the participant had some practice in the procedure before seeing a candidate sign, although they were not told this.

Recognition

After the participants had completed the legibility and meaning section of the test, the intended meaning of the crossover signs was explained to them and they were given prints of the seven signs to become familiar with them. The next section of the test concerned recognition of the signs once their meaning was known by the participants.

The instructions for Part III were read to the subjects and they were again shown the 20 slides, but in a different random order. The participants walked individually toward the projected sign until they could identify it. Participants were encouraged to guess the meanings of the signs as far away as possible from the screen so as to maximize confusion. All instances of confusion and the distance at which they occurred and the distance at which each sign was correctly identified were recorded.

When each sign had been correctly identified, the participant walked back to the 110-ft mark, the next slide was presented, and the procedure was repeated. This process was repeated until each participant had seen all 20 slides.

Preference

The last part of the experiment was a preference test. The participants were instructed to arrange prints of the seven crossover signs in order from the one they liked best to the one they liked least. The order in which the participants ranked each sign was then recorded. The participants were then asked seven questions about crossovers in general. A full description of the methodology is presented by Worsey *(4)*.

RESULTS

Legibility

Although the different experimental conditions preclude statistical comparison, the data in Table 1 indicate that the legibility distances for both groups of participants were similar. The distances for the Missouri participants were slightly longer in most but not all cases.

Understanding

The data in Table 2 indicate that the Missouri participants had more difficulty in guessing the meaning of the signs than the Virginia participants. This was particularly true for the symbol signs although only the arrows sign was guessed by more than one-half of the participants in Virginia. However, this difference was not statistically significant.

The total number of misinterpretations of the signs by uncued participants was approximately the same (98 in Virginia and 103 in Missouri). The data in Table 2 indicate that the Missouri participants generally misinterpreted the signs more often than the Virginia participants. They also failed more frequently to guess the meaning of the signs, with the exception of the permissive U-turn sign. However, these differences were not statistically significant.

In Virginia, the arrows sign was misinterpreted most often, followed by the nose plus arrows sign and the crossover sign. In Missouri, the nose plus arrows sign was misinterpreted most often, followed by the "crossover" sign and the permissive U-turn sign. The most frequent misinterpretations of the crossover signs were basically the same for both groups of participants.

The data in Table 2 indicate that the mean distances at which participants in both states understood the meaning of the signs were similar. The word signs were understood at much shorter distances, and of the symbol signs, the arrows sign was understood at the farthest distance.

Recognition

The data in Table 3 indicate that the mean distances at which participants in both states recognized the signs were somewhat similar, with the Missouri participants recognition distances being slightly shorter for all the signs except the arrows sign. The greatest difference was for the "median opening" sign, which Missouri participants recognized at a mean distance approximately 12 ft shorter than the distance Virginia participants recognized it. In both sets of results the Virginia crossover sign was recognized at by far the greatest average distance and the worded signs were recognized at the shortest distances.

The total number of instances in which participants confused the crossover signs with other signs was 20 for both data sets. These confusions followed a similar pattern for both data sets. In Missouri all the signs were recognized by all the participants, whereas in Virginia one participant did not recognize the crossover nose sign.

TABLE 1 COMPARISON OF SIGN FEATURE MEAN LEGIBILITY DISTANCES (ft)

		Type of Sign						
Feature	State	1 Crossover	2 Median Opening	3 Crossover Virginia	4 Crossover Nose	5 Crossover Nose Plus Arrows	6 Crossover Arrows	7 Permissive U-Turn
Sign shape	Virginia	100	107	63	102	104	108	101
	Missouri	104	106	65	102	103	109	102
Sign color	Virginia	106	108	99	106	107	108	100
	Missouri	109	108	96	108	108	105	86
Symbol or letter color	Virginia	75	71	57	79	84	89	55
	Missouri	83	76	55	93	95	101	76
Symbol or letter presence	Virginia	48	54	52	83	85	90	66
	Missouri	41	50	69	93	95	101	85
Median nose presence	Virginia	N/A	N/A	34	36	35	N/A	26
	Missouri	N/A	N/A	37	40	35	N/A	29
Road pattern	Virginia	N/A	N/A	N/A	N/A	34	52	25
	Missouri	N/A	N/A	N/A	N/A	38	59	28
Crossover movement	Virginia	N/A	N/A	N/A	N/A	N/A	48	25
	Missouri	N/A	N/A	N/A	N/A	N/A	50	27
Read legend	Virginia	12	11	N/A	N/A	N/A	N/A	N/A
	Missouri	12	12	N/A	N/A	N/A	N/A	N/A

NOTE: N/A = not applicable.

TABLE 2 COMPARISON OF THE UNDERSTANDING OF SIGNS IN VIRGINIA AND MISSOURI

	State	Type of Sign 1 Crossover	2 Median Opening	3 Crossover Virginia	4 Crossover Nose	5 Crossover Nose Plus Arrows	6 Crossover Arrows	7 Permis-sive U-Turn
Understanding distance (ft)	Virginia	14	11	25	33	31	41	22
	Missouri	13	12	26	29	28	37	19
Correct answer first attempt (freq)	Virginia	16	27	13	18	15	13	9
	Missouri	12	24	12	9	11	14	9
Incorrect guess before correct answer (freq)	Virginia	10	3	4	5	8	11	6
	Missouri	13	5	2	4	4	8	5
Did not make a correct guess (freq)	Virginia	4	0	13	7	7	6	15
	Missouri	5	1	16	17	15	81	16
Misinterpretations (freq)	Virginia	18	5	10	11	18	23	13
	Missouri	20	7	8	15	24	12	17
Subjects who would not attempt to guess meaning (freq)	Virginia	0	0	8	3	2	0	10
	Missouri	2	0	10	9	4	1	8

Preference

The data in Table 3 indicate that the Virginia participants had a much more clearly defined set of preferences than the Missouri participants. (Their mean preference rankings ranged from 3.07 to 6.00, whereas the Missouri participants' mean preference rankings had much less spread, from 3.52 to 4.76.)

The permissive U-turn sign, followed by the Virginia crossover sign, was least preferred by Virginia participants, whereas the Virginia crossover sign, followed by the permissive U-turn sign, was least preferred by Missouri participants. The Virginia participants most preferred the crossover nose sign, followed by the nose plus arrows sign and then the word signs. The Missouri participants most preferred the "median opening" sign, followed by the crossover nose sign and then the arrows and nose plus arrows signs. The rankings given to each sign by the Virginia and Missouri participants were found to be significantly different for all the signs except for "crossover" and the Virginia crossover signs.

All participants were asked their opinions on median crossovers; 80 percent of the Missouri participants considered crossovers to be hazardous whereas only 73 percent of Virginia participants considered them hazardous (Table 4). The types of hazards participants associated with median crossovers were slightly different for both groups. Traffic accelerating into the fast lane was considered as much of a hazard as traffic slowing in the fast lane by Missouri participants but not Virginia participants (Table 5). Missouri participants appeared to be more concerned with traffic crossing the divided highway than the Virginia participants (13 participants, compared with 4 Virginia subjects, mentioned traffic pulling out in front of them or turning traffic). One Missouri participant mentioned gravel crossovers as being dangerous.

The responses to the question, "What effect would a

TABLE 3 COMPARISON OF THE RECOGNITION AND PREFERENCE RANKINGS OF SIGNS IN VIRGINIA AND MISSOURI

	State	Type of Sign 1 Crossover	2 Median Opening	3 Crossover Virginia	4 Crossover Nose	5 Crossover Nose Plus Arrows	6 Crossover Arrows	7 Permis-sive U-Turn
Recognition distance (ft)	Virginia	39	42	82	48	47	57	61
	Missouri	34	29	76	48	45	60	58
Confusions (freq)	Virginia	2	2	0	2	8	5	1
	Missouri	0	2	0	3	9	6	0
Subjects who did not know the meaning (freq)	Virginia	0	0	0	1	0	0	0
	Missouri	0	0	0	0	0	0	0
Mean preferences (rank)	Virginia	3.47	3.37	4.57	3.07	3.23	4.30	6.00
	Missouri	4.10	3.52	4.76	3.62	3.79	3.66	4.62

TABLE 4 SUBJECTS' OPINIONS ON MEDIAN CROSSOVERS

	Yes				No			
	Virginia	Percent	Missouri	Percent	Virginia	Percent	Missouri	Percent
Do you think median crossovers constitute a hazard on a divided highway?	22	73	24	80	8	27	6	20
Do you think a sign would help identify a crossover if you wanted to use one?	29	97	29	97	1	3	1	3
Would the addition of a distance plate help you locate a crossover?	28	93	25	83	2	7	5	17

TABLE 5 TYPES OF HAZARDS ASSOCIATED WITH MEDIAN CROSSOVERS

	Frequency	
Hazard	Virginia	Missouri
Traffic slowing in fast lane	20	10
Traffic accelerating into fast lane	8	10
Turning traffic	4	8
Sudden lane changes	3	2
Traffic pulling out in front	0	5
Rear-end collisions	7	6
Broadside collisions	4	5
None	2	5

crossover sign have on your driving?" were basically the same for both sets of participants (Table 6). However, a larger number of the Missouri participants indicated that they would slow down if they saw a crossover sign (11 compared to 5 in Virginia) and that such a sign would have no effect on their driving (3 compared to 1 in Virginia).

When asked about word message signs, "crossover" was chosen by most participants in both groups (67 percent in Virginia and 76 percent in Missouri) as best conveying the intended meaning. "Opening" was the next-favored sign by the Virginia participants (23 percent) and "crossing" was the next favored by the Missouri participants (17 percent).

When asked the question, "Which word best conveys the presence of such a facility to you?" subjects responded as follows:

Crossover		*Crossing*		*Opening*	
Virginia (%)	*Missouri (%)*	*Virginia (%)*	*Missouri (%)*	*Virginia (%)*	*Missouri (%)*
20	23	3	5	7	2
67	76	10	17	23	7

When asked, "Would the addition of the word "median" help to clarify the meaning of the sign?" subjects responded as follows:

Yes		*No*	
Virginia (%)	*Missouri (%)*	*Virginia (%)*	*Missouri (%)*
24	22	6	8
80	73	20	27

The distances at which participants responded that they believed the sign should be placed in front of a crossover tended to be greater in Missouri than in Virginia. This is reflected in the mean distances, which were 838 ft in Virginia and 1,322 ft in Missouri.

Word Message Signs

A median crossover sign can be worded or symbolic. The "median opening" sign was the word message sign understood best by uncued participants in Virginia, and the majority of them chose "crossover" as conveying the intended meaning better than "crossing" or "opening." A "median crossover" sign was therefore made and shown to those Missouri participants for which there was time available to do so.

The data in Table 7 indicate that the legibility, understanding, and recognition distances for the "median crossover" sign were about the same as that for the other word message signs. An intermediate percentage of uncued participants guessed the meaning of the "median crossover" sign without a wrong guess first (87 percent compared with 90 percent of the Virginia participants and 80 percent of the Missouri participants for the "median opening" sign). All of the participants eventually managed to guess the meaning of the "median crossover" sign. There were only three misinterpretations of the "median crossover" sign by the uncued participants compared with five for the "median opening" sign in Virginia and seven in Missouri. There were no instances of confusion with other signs once the participants had had the meaning of the sign explained to them, whereas the "median opening" sign was confused with other signs twice in both Virginia and Missouri.

TABLE 6 EFFECT OF SIGN ON SUBJECTS' DRIVING

	Frequency	
Effect	Virginia	Missouri
Would look for sign if wanted to use a crossover	12	7
Would change lanes if wanted to use a crossover	1	1
Would signal if wanted to use a crossover	1	0
Would look for slowing traffic	15	12
Would slow down	5	11
Would change lanes	4	3
None	1	3

TABLE 7 COMPARISON OF THE "MEDIAN CROSSOVER" SIGN WITH THE "MEDIAN OPENING" AND "CROSSOVER" SIGNS

	Crossover		Median Opening		Median Crossover
	Virginia	Missouri	Virginia	Missouri	Missouri
Legibility distances (ft)					
Sign shape	100	104	107	106	102
Sign color	106	109	108	108	107
Legend color	75	83	71	76	93
Letter presence	48	41	54	50	49
Read legend	12	12	11	12	12
Understanding distance (ft)	14	13	11	12	12
Recognition distance (ft)	39	34	42	29	28
Correct answer	16	12	27	24	21
First attemtp (freq)					
Percent	53	40	90	80	87
Incorrect answer	10	13	3	5	3
Before correct one (freq)					
Percent	33	43	10	17	13
Don't know (freq)	4	5	0	1	0
Percent	13	17		3	
Misinterpretations (freq)	18	20	5	7	3
Confusions (freq)	2	0	2	2	0

CONCLUSIONS

From this study there appears to be justification for the use of signs indicating the presence of a median crossover that can be used by the general public. The majority of participants tested in both groups perceived crossovers as hazardous locations, and from their responses to the questions they were clearly aware of the potential hazards that crossovers can cause. Most participants indicated that if such a sign were installed, it would likely have a beneficial effect on their driving behavior.

Although word message signs can usually be understood once they are read, they are not as legible as symbolic signs. Of the symbolic signs, the arrows sign had the best average legibility and understanding distances in both Virginia and Missouri, but it had by far the most misinterpretations by participants in Virginia. Although it was ranked second among the symbol signs by the Missouri participants in terms of preference, it was ranked fifth by the Virginia participants and is therefore not recommended.

Of the other symbolic signs, the permissive U-turn sign had low average legibility and understanding distances in both Virginia and Missouri and was not well understood by the participants. This is reflected in its being ranked last in the preference test by the majority of participants in Virginia and many in Missouri. The significance of the green periphery to indicate a permissive sign was not understood, and this sign is not recommended.

Of the symbolic signs, the Virginia crossover sign also had low average legibility distances and again was not well understood by uncued participants in both Virginia and Missouri. In the preference test it was not well liked by either group of participants. However, it did very well in the recognition test in both Virginia and Missouri, presumably because of its different color and shape. It was recognized at a far greater average distance than any of the other signs and was the only sign not confused in Virginia. Several participants in both Virginia and Missouri mentioned that if they had initially known the meaning of the sign they believed it would be the best one to use. The meaning of the sign was not obvious to the participants in either Virginia or Missouri. However, in Virginia the sign is placed at the median opening, which should lead to a high degree of self-education.

Of the remaining symbolic signs, the nose plus arrows sign had slightly better average legibility distances but the crossover nose sign had slightly better average understanding and recognition distances in both Virginia and Missouri. The latter sign also had fewer misinterpretations and instances of confusion in the understanding and recognition sections of the experiment than the former in both Virginia and Missouri. It was also given the best average rank out of all the signs in the Virginia preference test and the best average rank out of the symbol signs in Missouri. It also had the simplest design of all the signs tested. Of the symbol signs tested, the crossover nose sign (see Figure 2) is recommended to indicate the presence of a median crossover.

Despite the different experimental conditions, the legibility, understanding, and recognition distances of all the signs were similar for both groups of participants. However, the Missouri participants had more difficulty identifying the green (Virginia crossover and permissive U-turn) signs than the Virginia participants.

The Missouri participants had more difficulty than the Virginia participants in guessing the meaning of nearly all the signs, especially the symbol signs. They misinterpreted the signs more often and could not guess the meaning of the signs as frequently.

The greatest differences between the Virginia and Missouri results were in the preference rankings the participants gave to the signs. The Virginia participants had a much more clearly defined set of preferences, whereas the Missouri participants' preferences were much more evenly spread with little agreement among the participants. The Missouri participants also

TABLE 8 SUMMARY OF FAVORABLE AND UNFAVORABLE FINDINGS FOR EACH CROSSOVER SIGN

Sign Type	Favorable Aspects	Unfavorable Aspects
Word message	Usually understood once read	Much less legible than symbolic signs
Arrows	Most legible sign Understood the farthest away	Misinterpreted the most Not liked by subjects
Permissive U-turn	Not confused with other signs	Legible at shorter distances Understood the closest out of symbol signs Least liked by most subjects Meaning of green periphery not understood
Virginia crossover	Recognized the farthest away Not confused with other signs	Legible at shorter distances Understood the second closest out of symbol signs Not liked by subjects Not understood well
Nose plus arrows	Second most legible of symbol signs	
Crossover nose	Understood the second farthest away Misinterpreted the second least of symbol signs Most preferred of symbol signs	

preferred the word message signs more than the Virginia participants. This was especially true for females, particularly those over 50. The Virginia participants preferred the crossover nose and nose plus arrows signs over the word message signs, whereas the Missouri participants preferred the "median opening" sign.

Although there were some differences in the Virginia and Missouri results, the same conclusions were reached—that a "median crossover" sign would be the best word message sign to use and the crossover nose sign would be the best symbolic sign to use to indicate the presence of a median crossover. Despite the Missouri participants' preferences for word message signs, legibility of the symbolic signs was so much greater that the crossover nose sign is the sign recommended for field evaluation to identify median crossovers. Table 8 contains a summary of the findings pertaining to the signs tested in this study and the impressions and preferences expressed by the subjects.

ACKNOWLEDGMENT

A substantial portion of this study was supported by a Grant for Research Fellowship awarded by the Federal Highway Administration, U.S. Department of Transportation. Special thanks are extended to the following individuals and organizations for their guidance, suggestions, and technical help: Samuel Tignor and Howard Bissell of the Federal Highway Administration and R. N. Robertson and C. Morris of the Virginia Department of Highways and Transportation. The opinions, interpretations, and conclusions expressed in this paper are solely those of the authors.

REFERENCES

1. P. D. Cribbins, J. M. Arey, and J. K. Donaldson. Effects of Selected Roadway and Operational Characteristics on Accidents on Multi-Lane Highways. In *Highway Research Record 188*, HRB, National Research Council, Washington, D.C., 1967, pp. 8–25.
2. *Manual on Uniform Traffic Control Devices, Revision 3*. FHWA, U.S. Department of Transportation, 1978.
3. J. Walker, E. Alicandri, and K. Roberts. *An Evaluation of Candidate Symbolic Routing Signs for Trucks Carrying Hazardous Cargo*. Report FHWA-RD-85-081. FHWA, U.S. Department of Transportation, 1985.
4. G. M. Worsey. *An Investigation of Signs for Median Crossovers*. Ph.D. dissertation. University of Missouri-Rolla, 1986.

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

Abridgment

Effects of Reduced Speed Limits in Rapidly Developing Urban Fringe Areas

GERALD L. ULLMAN AND CONRAD L. DUDEK

Speed zoning on the basis of the 85th percentile speed in rapidly developing urban fringe areas usually results in the posting of 55 mph speed limits. Although these areas have some urban-like characteristics, no differentiation in speed limits is made between highways in these areas and those in rural locations. Speed zoning below the 85th percentile may be beneficial to drivers in rapidly developing areas, indicating that the area requires additional attention and caution. Presented in this paper are the results of field studies conducted at six urban fringe highway sites in Texas where speed limits were currently 55 mph and rapid urban development was occurring. Speed zones of 45 mph were installed at these sites even though the 85th percentile speed did not warrant the lower speed zones. Spot speed, speed profile, and accident data were collected before and after the speed zones were implemented. No significant changes occurred in speeds, speed distributions, or speed-changing activity at the sites. Likewise, accident rates remained unchanged. It appears that the lower speed zones were not effective in improving safety at these sites.

In recent years, several sections of highways on the fringes of many major cities in Texas have been experiencing rapid urban development. The driving environment on these highways has become more complex as traffic volumes increased, adjacent commercial and residential dwelling units were constructed, and new and additional forms of traffic control were installed during a short period of time. At many locations, accidents and accident rates have increased significantly. As a result of the high speeds still present on these highways, many of the accidents have been quite severe.

Current speed zoning procedures *(1)*, which rely primarily on the 85th percentile speed of traffic on a facility, may not be adequate for these rapidly developing urban fringe areas. Even though the areas develop some urban characteristics, the 85th percentile speed usually indicates that a speed limit of 55 mph be posted, identical to that posted in rural areas. In effect, no distinction in speed limits is made between highway sections in a rural area and highway sections undergoing rapid development in urban fringe areas.

Speed zoning procedures might be improved by allowing a speed limit to be posted below the 85th percentile speed in these rapidly developing urban fringe areas. This action may signal to motorists that the driving environment is more complex and that additional attention and caution is needed. To test this hypothesis, the Texas Transportation Institute has conducted a study sponsored by the Texas State Department of Highways and Public Transportation to examine the effect of implementing speed limits below the 85th percentile speed on highways in rapidly developing urban fringe areas.

Texas Transportation Institute, The Texas A&M University System, College Station, Tex. 77840.

STUDY METHOD

Before-and-after speed and accident data were collected to evaluate the effectiveness of implementing speed zones below the 85th percentile speed. Six study sites on two-lane and four-lane undivided highways were identified where (a) rapid urban development was occurring in urban fringe areas that had been primarily rural in nature, and (b) 55 mph speed limits were still posted. Characteristics of study sites are given in Table 1. Speed and accident data were collected at each site after which the speed limits were reduced to 45 mph. Speed and accident data were collected again after the speed limits were lowered. The two sets of data were then analyzed and compared.

Speed limits of 45 mph were selected for study because it was believed that 50 mph limits would not present the same sense of urbanization to motorists, whereas zones of 40 mph or below would be too inconsistent with existing speeds on the facility and would immediately be dismissed by drivers as unreasonable and unrealistic. Also, the study was designed to investigate only the effects of reduced speed limits on traffic speeds and accidents. Consequently, attempts were made to maintain law enforcement at a constant level during the study, and no public notice was given of the speed limit reductions. Although the effects of additional enforcement or public notification of the speed limit reduction would have been of interest, determining them was beyond the scope of this study.

Data Collection and Reduction

Spot speed data were collected at three locations placed one-fourth, one-half, and three-fourths of the way through each study site. In each direction at each location, speeds of 125 free-flowing vehicles were obtained by using a speed radar gun from within a vehicle parking in as inconspicuous a position as possible. Free-flow vehicles were defined as vehicles having at least a 5 sec headway between them and the vehicle directly in front of them. Consequently, both isolated vehicles (those with no other vehicles nearby) and vehicles at the head of platoons were eligible for sampling. An attempt was made to sample isolated vehicles and lead vehicles in platoons in proportion to their relative frequencies at the sites. At the four-lane study sites, an attempt was also made to sample from each lane in

TABLE 1 SUMMARY OF STUDY SITE CHARACTERISTICS

Site	Location	Length (mi.)	Cross-Section	Development Degree	Development Type	1983 AADT	Accidents/MVM 1983
1	Houston	2.3	2-lane, 2-way	Low	Residential	14,100	3.9
2	Houston	3.1	2-lane, 2-way	Low	Residential	10,700	1.2
3	Houston	3.9	4-lane, Undivided	Moderate	Residential	30,400	0.9
4	Houston	3.3	4-lane, Undivided	High	Commercial, Residential	29,000	4.7
5	Austin	2.0	4-lane, Undivided	Moderate	Commercial, Residential	25,000	8.2
6	Ft. Worth	2.3	4-lane, Undivided	Low	Commercial, Residential	11,500	3.7

AADT = Annual Average Daily Traffic

MVM = Million-Vehicle-Miles

proportion to the volumes of traffic on each. Care was taken not to choose data collection locations near intersections, major driveways, or other features that may affect normal driving speeds. Before-and-after data were collected at the same locations to ensure that locational differences did not affect the study results. Several statistics of interest were computed from the spot speed data including

1. Average speed,
2. The 85th percentile speed,
3. Proportion of recorded speeds exceeding 60 mph,
4. Standard deviation of speeds, and
5. Skewness index of the distribution of speeds.

Selection of these statistics were based on documented relationships between vehicle speeds and accident severity and frequency. Because accident severity appears directly linked to vehicle speed *(2, 3)*, the average speeds, 85th percentile speeds, and the proportion of vehicles exceeding 60 mph were used to measure overall changes in speeds and the number of high-speed drivers at the study sites. In contrast to accident severity, accident frequency appears less dependent on absolute speeds than on either the variability of speeds *(4)* or the shape of the speed distribution *(5, 6)*. Consequently, the standard deviation of speeds, as well as the skewness index of the speed distribution, were computed from the data collected.

In addition to spot speeds, speed profile data were also collected at the sites. A car-following technique was used with an instrumented vehicle to obtain measurements of speed every 500 ft through the sites. Twenty vehicles selected at random were followed through each direction of travel at each site. A measure of speed-changing activity was computed from the profile data, based on the acceleration noise concept originally introduced by Jones and Potts *(7)*, and successfully used by others *(8)* for describing the quality of traffic flow in quantitative terms. Acceleration noise is defined as the standard deviation of the accelerations and decelerations of an individual vehicle as it travels over a particular section of road. This value represents the disturbance of the vehicle's speed from a uniform speed and provides a measure of the frequency and degree of speed changes for that vehicle.

Accident data from the Master Accident File maintained by the Texas Department of Public Safety were obtained for the 1-year period before installation of the 45 mph speed zones at the study sites. The zones were left in place for 1 year, at which time the accident data for that year were also obtained from the Master Accident File. Because the study sites were located in urban fringe areas of ongoing development, significant changes in traffic volumes occurred over the 2-year study period, as observed from Table 2. Therefore, accident rates (accidents per million vehicle miles) were computed at each site for both total and severe (fatal and injury) accidents.

TABLE 2 CHANGES IN TRAFFIC VOLUMES

Site	Traffic Volumes (AADT) Before	After	Change
1	14,000	16,200	+16%
2	10,300	10,600	+3%
3	33,000	35,000	+6%
4	22,000	19,800	-10%
5	27,000	31,000	+15%
6	16,800	18,000	+7%

One aspect of speed control proven to have a dramatic effect on vehicle speed is the degree of law enforcement *(9, 10)* at a location. The law enforcement agencies responsible for patrolling and speed-ticketing at the various study sites were requested to maintain the same level of enforcement efforts throughout the study. Although all agencies did agree to maintain their current efforts, objective data (such as the number of speeding tickets given during the before and after time periods) were not available to check to determine if enforcement levels remained constant. Also, this lack of information prevented a comparison of relative levels of enforcement between sites.

RESULTS

Effect on Speeds

Overall, the installation of 45 mph speed limits at the study sites in rapidly developing urban fringe areas had little effect on vehicle speeds. A summary of the average speeds, 85th percentile speeds, and the proportion of drivers exceeding 60 mph are given in Table 3. These results are for the middle data collection location at the study sites. Results for the other two locations where spot speed data were taken are included in the Appendix. Although slight location-to-location variation did exist at the sites, the overall changes between the before-and-after speed data were similar at all locations. As can be seen in Table 3, Site 5 experienced a 4 to 6 mph reduction in the average and 85th percentile speeds. Likewise, the proportion of drivers exceeding 60 mph dropped 6 to 10 percent. However, the remaining sites did not experience similar reductions in speeds; speeds actually rose slightly at Site 3. It does not appear that the lower speed zones were consistently effective in reducing vehicle speeds at these sites to any significant degree. It is possible that the reduction in speeds at Site 5 were the result of a relatively higher level of law enforcement as compared to the other sites, but this fact is not known for certain because the information about law enforcement efforts at the sites was not available.

Examination of the standard deviation, skewness index, and acceleration noise statistics also suggests that the lower speed zones had little or no effect on the speed distribution or speed-changing activity at the study sites. The data in Table 4 illustrate how these statistics generally did not change between the before-and-after speed data collected at the sites.

Effect on Accidents

A comparison of accident rates at the six sites is given in Table 5. Although two sites did experience a reduction in accidents, the overall evaluation generally showed no change in accident

TABLE 3 EFFECT OF 45 mph SPEED ZONES ON VEHICLE SPEEDS

Site	Average Speed (mph)			85th Percentile Speed (mph)			Proportion of Drivers Exceeding 60 mph (%)		
	Before	After	Change	Before	After	Change	Before	After	Change
1									
EB	47.3	47.0	-0.3	53	52	-1	0	0	0
WB	47.8	48.3	+0.5	54	53	-1	1	8	+7*
2									
EB	53.2	52.3	-0.9	61	58	-3	21	9	-12*
WB	53.2	52.8	-0.4	59	59	0	14	15	+1
3									
EB	48.5	52.3	+3.8*	59	57	+3	1	6	+5*
WB	49.2	49.9	+0.7	54	54	0	6	3	-3
4									
NB	42.9	43.4	+0.5	49	49	0	1	15	+14*
SB	44.8	43.6	-1.2	50	48	-2	1	12	+11*
5									
NB	53.1	47.2	-6.1*	58	53	-5	11	1	-10*
SB	51.1	46.9	-4.2*	56	52	-4	7	1	-6*
6									
NB	52.9	51.9	-1.0	59	57	-2	11	9	-2
SB	54.2	49.9	-3.3*	59	56	-3	12	6	-6

*Statistically Significant Change from Before Condition (Level of Confidence = 95%)

TABLE 4 EFFECT OF 45 mph SPEED ZONES ON THE DISTRIBUTION OF SPEEDS AND SPEED-CHANGING ACTIVITY

Site	Standard Deviation (mph) Before	After	Change	Skewness Index[a] Before	After	Acceleration Noise (ft/sec^2) Before	After	Change
1								
EB	5.2	5.1	-0.1	1.0	1.0	1.1	1.0	-0.1
WB	5.2	4.9	-0.3	1.0	1.0	1.3	1.0	-0.3*
2								
EB	7.3	6.4	-0.9	0.9	1.0	-0.9	0.9	0.0
WB	5.6	6.1	+0.5	0.9	1.1	0.7	0.6	-0.1
3								
EB	4.8	4.7	-0.1	0.9	0.9	1.2	0.9	-0.3*
WB	5.8	5.1	-0.7	1.1	0.8	1.3	1.1	-0.2*
4								
NB	6.3	6.3	0.0	1.0	0.9	1.1	1.2	+0.1
SB	5.3	6.0	+0.7	1.0	0.9	1.1	1.1	0.0
5								
NB	5.9	5.5	-0.4	0.9	0.9	1.1	1.3	+0.2
SB	6.0	5.3	-0.7	0.9	0.9	1.3	0.9	-0.3*
6								
NB	5.8	6.1	+0.3	1.1	1.0	1.0	0.9	-0.1
SB	4.8	5.6	+0.8	0.9	0.9	0.7	0.8	+0.1

*Statistically Significant Change from Before Condition (Level of Confidence = 95%)

[a]Skewness Index was computed as $\frac{2(\text{93rd \%-tile} - \text{50th \%-tile speeds})}{(\text{93rd \%-tile} - \text{7th \%-tile speeds})}$

rates. Similarly, severe accidents did not appear to be reduced; Sites 2 and 6 posted increases in the frequency of these accidents. Only Site 3 was found to have a substantial reduction in its severe accident rate. Curiously, this was also the site at which speeds increased after the 45 mph speed zones were installed. As only 1 year of after accident data were available, and because accidents in themselves are rare events, the changes in accident rates that occurred were most likely the result of random fluctuation regression-to-the-mean, rather than a reduction in the posted speed limit.

SUMMARY OF FINDINGS

The effect of reducing speed limits below the 85th percentile speed of traffic at locations in rapidly developing urban fringe areas has been examined in this study. Overall, a reduction in the speed limits from 55 to 45 mph at the six study sites had no conclusive effect on absolute speeds, speed distributions, or speed-changing activity. Likewise, the lower limits were not effective in reducing the frequency or the severity of accidents occurring at the study sites. It is not known whether motorists did not notice the reduced speed limits, or whether drivers saw but chose to disregard or ignore the lower limits. Whatever the reason, the lower speed limits in rapidly developing urban fringe areas did not persuade motorists to drive more carefully.

These results parallel those of past research efforts *(11)* that have attempted to influence drivers to operate their vehicles at a "safer" speed by posting lower speed limits. As this and previous studies have revealed, reduced speed limits apparently do not alter, to any significant degree, perceptions of accident risk, the potential of receiving a speeding ticket, or any of the other factors that drivers are assumed to consider when selecting the speed at which they travel.

The study results show that traffic safety and operations were not improved in rapidly developing urban fringe areas solely by posting a speed limit below the 85th percentile speed. As stated previously, this study was unable to consider the effects of

TABLE 5 COMPARISON OF ACCIDENT RATES

Site	Rate Before	Rate After	Change (%)
Total Accidents			
1	4.08	2.57	−37[a]
2	1.11	1.08	−3
3	2.02	1.22	−40[a]
4	7.32	9.14	+25
5	7.10	7.03	−1
6	2.41	3.04	+26
Severe (Fatal and Injury) Accidents			
1	1.53	1.47	−4
2	0.26	0.58	+125
3	0.83	0.46	−44
4	2.98	2.98	NR
5	3.15	2.79	−11
6	0.92	1.66	+80[a]

NOTE: Accidents per million vehicle-miles. NC = No change in accident rates.

[a]Significant change in accident rate based on Poisson comparison of means test (level of confidence = 95 percent).

increased law enforcement or public notification at the speed limit reduction, or both, on speed and accidents. Additional research to examine the effect of these factors should be considered, with special emphasis on whether the costs of implementing these factors are justified through reduced accident costs and improved traffic operations on highways in rapidly developing urban fringe areas.

REFERENCES

1. *Procedures for Establishing Speed Zones.* Texas State Department of Highways and Public Transportation, Austin, 1985.
2. D. Solomon. *Accidents on Main Rural Highways Related to Speed, Driver, and Vehicles.* Bureau of Public Roads, 1964.
3. H. C. Joksch. An Empirical Relation Between Fatal Accident Involvement and Speed. *Accident Analysis and Prevention,* Vol. 7, 1975.
4. R. Michaels and C. Schnieder. The Energy Crisis, Characteristics of Traffic Flows and Highway Safety. *Accident Analysis and Prevention,* Vol. 8, Dec. 1976.
5. W. C. Taylor. Speed Zoning: Theory and Its Proof. *Traffic Engineering,* Vol. 35, No. 4, Jan. 1965.
6. R. J. Krazemski. *The Use of Skewed Speed Distributions to Locate Points of High Accident Potential on Low-Volume Two-Lane Rural Highways.* Report JHRP-76-22. Purdue University, Lafayette, Ind., Aug. 1976.
7. T. R. Jones and R. B. Potts. The Measurements Acceleration Noise—A Traffic Parameter. *Operations Research,* Vol. 10, No. 6, 1962.
8. D. R. Drew, C. L. Dudek, and C. J. Keese. Freeway Level of Service as Described by an Energy-Acceleration Noise Model. In *Highway Research Record 162,* HRB, National Research Council, Washington, D.C., 1967, pp. 30–85.
9. K. B. Joscelyn, T. H. Bryan, and D. M. Goldenbaum. A Study of the Effects of Law Enforcement on Traffic Flow Behavior. Report DOT-HS-800-505. NHTSA, U.S. Department of Transportation, Jan. 1971.
10. O. K. Dart, Jr. and W. W. Hunter. Evaluation of the Halo Effect in Speed Detection and Enforcement. In *Transportation Research Record 609,* TRB, National Research Council, Washington, D.C., 1976, pp. 31–35.
11. D. Warren. Chapter 17, Speed Zoning and Control. In *Synthesis of Safety Research Related to Traffic Control and Roadway Elements.* Report FHWA-TS-82-233. FHWA, U.S. Department of Transportation, Dec. 1982.

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

APPENDIX

TABLE A-1 EFFECT OF 45 mph SPEED ZONES ON VEHICLE SPEEDS, LOCATION 1

Site	Average Speed (mph) Before	After	Change	85th Percentile Speed (mph) Before	After	Change	Proportion of Drivers Exceeding 60 mph (%) Before	After	Change
1									
EB	47.7	45.9	-1.8*	51	52	+1	2	1	-1
WB	49.1	45.5	-3.6*	54	52	-2	3	0	-3
2									
EB	55.2	53.8	-1.4	61	60	-1	24	20	-4
WB	55.1	54.1	-1.0	60	60	0	18	17	-1
3									
EB	46.8	48.9	+2.1**	53	53	0	1	2	+1
WB	47.6	50.0	+2.5**	53	56	+3	3	5	+2
4									
NB	53.2	50.2	-3.0	59	56	-3	15	9	-6*
SB	52.3	50.5	-1.8	58	58	0	11	9	-2
5									
NB	51.5	48.3	-3.2*	57	53	-4	9	2	-7*
SB	51.0	46.8	-4.2*	56	52	-4	6	0	-6*
6									
NB	49.1	47.3	-1.8*	56	52	-4	3	3	-1
SB	49.7	47.4	-2.2*	55	54	-1	3	1	-2

**Statistically Significant Increase from Before Condition (Level of Confidence = 95%)

*Statistically Significant Decrease from Before Condition (Level of Confidence = 95%)

TABLE A-2 EFFECT OF 45 mph SPEED ZONES ON THE DISTRIBUTION OF SPEEDS, LOCATION 1

Site	Standard Deviation (mph) Before	After	Change	Skewness Index Before	After
1					
EB	6.2	4.8	-1.4*	0.9	0.9
WB	6.3	5.0	-1.3*	1.1	1.0
2					
EB	5.9	6.2	+0.3	0.8	0.7
WB	5.4	6.0	+0.6	1.1	1.0
3					
EB	6.2	4.3	-1.9*	1.0	0.8
WB	6.1	5.6	-0.5	0.9	1.1
4					
NB	5.7	5.7	0.0	1.1	1.2
SB	6.2	6.8	+0.6	0.8	1.0
5					
NB	5.7	5.2	-0.5	1.1	0.8
SB	6.0	5.5	-0.5	0.9	0.9
6					
NB	6.8	4.8	-2.0*	0.8	0.9
SB	6.0	6.0	0.0	0.9	1.0

**Statistically Significant Increase from Before Condition (Level of Confidence = 95%)

*Statistically Significant Decrease from Before Condition (Level of Confidence = 95%)

TABLE A-3 EFFECT OF 45 mph SPEED ZONES ON VEHICLE SPEEDS, LOCATION 3

Site	Average Speed (mph) Before	After	Change	85th Percentile Speed (mph) Before	After	Change	Proportion of Drivers Exceeding 60 mph (%) Before	After	Change
1									
EB	41.3	44.8	+3.5**	47	50	+3	0	0	0
WB	44.3	45.2	+0.9	50	50	0	0	0	0
2									
EB	52.7	50.0	-2.7*	58	56	-2	10	1	-9*
WB	50.2	51.6	+1.4	55	58	+3	8	10	2
3									
EB	53.5	51.9	-1.6*	59	57	-2	14	2	-12*
WB	52.0	54.1	+2.1**	57	60	+3	7	17	+10**
4									
NB	47.1	48.4	+1.3	54	54	0	2	2	0
SB	46.3	48.4	+2.0**	52	54	+2	3	6	+3
5									
NB	54.4	48.6	-5.8*	60	54	-6	16	2	-14*
SB	49.5	48.1	-1.4	55	54	-1	6	3	-3
6									
NB	50.4	48.3	-2.1*	55	53	-2	4	2	-2
SB	53.2	49.4	-3.9*	59	54	-5	14	5	-9*

**Statistically Significant Increase from Before Condition (Level of Confidence = 95%)

*Statistically Significant Decrease from Before Condition (Level of Confidence = 95%)

TABLE A-4 EFFECT OF 45 mph SPEED ZONES ON THE DISTRIBUTION OF SPEEDS, LOCATION 3

Site	Standard Deviation (mph) Before	After	Change	Skewness Index Before	After
1					
EB	5.1	5.3	+0.2	0.8	1.3
WB	5.7	4.6	-1.1*	1.0	0.9
2					
EB	6.3	5.1	-1.2*	0.8	1.1
WB	4.7	6.9	+2.2**	1.0	0.9
3					
EB	5.8	4.5	-1.3*	1.1	0.9
WB	5.7	6.2	+0.5	0.9	1.0
4					
NB	5.9	5.1	-0.8	1.1	1.1
SB	6.4	6.1	-0.3	0.8	1.0
5					
NB	5.4	5.7	+0.3	0.8	0.9
SB	6.2	6.3	+0.1	0.9	1.0
6					
NB	4.9	4.7	-0.2	0.9	0.9
SB	5.9	4.9	-1.0*	1.0	1.0

**Statistically Significant Increase from Before Condition (Level of Confidence = 95%)

*Statistically Significant Decrease from Before Condition (Level of Confidence = 95%)

Traffic Operations of Basic Traffic-Actuated Control Systems at Diamond Interchanges

CARROLL J. MESSER AND MYUNG-SOON CHANG

This paper contains the results of field studies conducted to evaluate four types of basic, full-traffic-actuated signal control systems operated at three diamond interchanges. Two signal phasing strategies were tested: (a) three-phase and (b) four-phase with two overlaps. Two small-loop (point) detection patterns—single- and multipoint—were evaluated for each type of phasing. An assessment of these systems was conducted based on the results of statistical and observational evidence regarding their operational effects on queues and cycle lengths. Multiple and geometric linear regression were used to formulate models that relate queuing delay to traffic characteristics. Single-point detection was found to be the more cost-effective three-phase design. Multipoint detection was found to be the more delay-effective four-phase configuration. Four-phase control characteristically operates at a longer cycle length than does three-phase for a given traffic volume, and this feature may produce higher average delays unless the cycle increase is controlled to the extent that the internal progression features of four-phase can overcome this deficiency.

Efficient diamond interchange traffic control is a desirable objective and a necessary condition for providing safe and economic urban mobility. The diamond interchange is a critical interface between the freeway and arterial street system and, potentially, a system-threatening bottleneck to efficient traffic flow in an urban area.

Diamond interchanges are widely used in urban areas as a means to transfer freeway traffic to and from the surface street system. These interchanges are almost always signalized with traffic-actuated or pretimed signals *(1–4)*. This subject is addressed in this paper and useful information is provided for guiding future engineering decisions regarding the design and operation of traffic-actuated signals at diamond interchanges.

OBJECTIVES

This paper contains the results of field studies conducted to evaluate four types of basic, full traffic-actuated signal control systems. Two signal phasing strategies were tested: (a) three-phase, and (b) four-phase with two overlaps. Two small-loop (point) detection patterns, single- and multipoint, were evaluated for each type of phasing. An assessment of these four systems was conducted based on the results of statistical and observational evidence regarding their operational effects on queues and cycle lengths. Multiple and geometric linear regression were used to formulate models that relate queuing delay to traffic characteristics.

Description of the four control systems will be provided by the two principal categories of control; namely, three-phase and four-phase. All signal control systems tested provided basic, full-actuated control. No volume-density features were permitted. All systems were tested at diamond interchanges having continuous one-way frontage roads rather than exit ramps.

It was desired that the signal control units would be equally fine-tuned in the field by experienced traffic engineers to provide reasonably snappy operations. Gap sizes and minimum greens were set reasonably short for the various detector designs. No tendency to prematurely gap out within starting platoons was observed. In all cases, the same maximum phase settings (Max 1 and Max 2) were applied to the three-phase and four-phase control strategies. In retrospect, however, it cannot be proved that the actuated systems were equally fine-tuned, as no metric exists for this purpose. Therefore, direct comparisons between the operational performance of three-phase and four-phase control, in particular, should reflect the limitation of this study.

THREE-PHASE CONTROL

Phasing

The basic three-phase system used for traffic-actuated control of diamond interchanges in Texas is shown in Figure 1. Although there are three primary phases, six subordinate phases also are possible, depending on phase gap-out, phase calls, and controller programming, including ring rotation and overlaps.

Phase 1 initiates the sequence and includes both frontage road green signals to simultaneously provide protected movement into the interchange. This phase must be displayed if there is a call for either frontage road green. Following Phase 1, an extension of one of the two frontage road phases usually occurs during peak hours of traffic demand. The selected extension phase would reflect which green had the higher ramp volume.

Phase 2 is the cross-street, inbound-outbound phase without protected left turns. Inbound traffic is entering the interchange; outbound is exiting. Permissive left turns are sometimes allowed in Phase 2. Phase rotation from Phase 3 back to Phase 2 may occur during light traffic conditions when no frontage road

C. J. Messer, Texas A&M University, College Station, Tex. 77843. M. S. Chang, Korea Highway Traffic Safety Association, P.O. Box 376, KWHAG WGA MOON, Seoul, Korea.

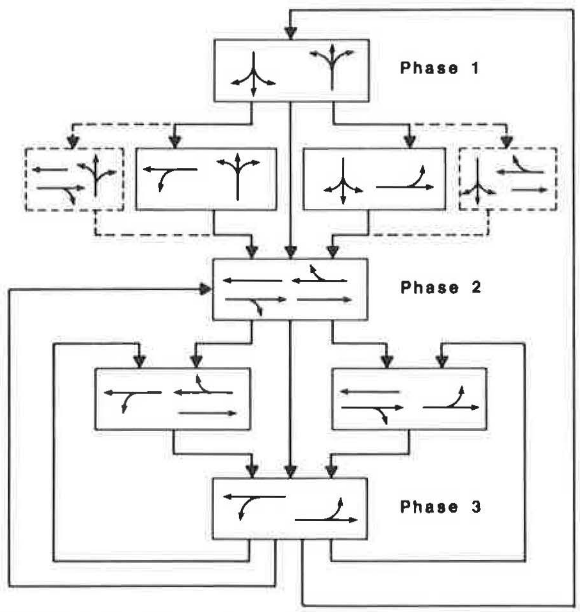


FIGURE 1 Three-phase, full-traffic-actuated diamond interchange phasing.

calls exist. Gap-out of an inbound through movement results in an early protected left-turn phase occurring before Phase 3.

Phase 3 is the simultaneous display of protected turn signals for both internal left turns serving outbound traffic. Both turn signals must simultaneously terminate. Right-of-way then normally goes to Phase 1 to start the sequence again.

Detectors

Two types of detector configurations were studied for three-phase control: (a) single-point and (b) multipoint. Similar designations were also given to detector configurations for four-phase control. However, as subsequent coverage will show, the detector configurations for four-phase control on the frontage roads were considerably different for both cases.

Single-point detectorization for three-phase control provides a minimal number of detectors at the interchange while still maintaining full-actuated control; that is, at least one detector station per approach. Although there were on-site variations because of approach speed, geometry, and the presence or absence of left-turn bays, a basic plan existed for each detector configuration. In the single-point detector plan, one detector was placed on each frontage road approach. Detector setback from the stop bar varied with approach speed, but often was about 100 ft. This placement provides a minimum required phase time of about 14 sec. Phase operations are concurrent for the two frontage roads with memory "on" (locking memory). Detector placement for the single loop sensor per cross-street inbound approach again depends on the approach speed, but averaged only about 100 ft to the stop bar.

Multipoint detection in three-phase control added one more detector across all lanes on all inbound phases. One detector was located about 100 ft from the stop bar as in single-point detection, and the other detector was located midway to the stop bar at about 50 ft. Again, actual detector placement depends on approach speed.

Figure 2 shows explicitly the three-phase multipoint detection scheme. Single-point detection did not include the inner approach detectors. Multipoint detection permits a slightly smaller minimum green with only slightly smaller gaps for extension timing.

FOUR-PHASE CONTROL

Phasing

This type of signal phasing provides four primary input phases to the interchange, with additional input capacity provided by judicious arrangement of the four basic phases to allow two adjustable, fixed-duration overlap phases. This signal strategy is commonly referred to as "four-phase with overlaps." In reality, six discrete phases are required when all phases are calling. The phasing sequence is shown in Figure 3. Note that phase numbering is different between three-phase and four-

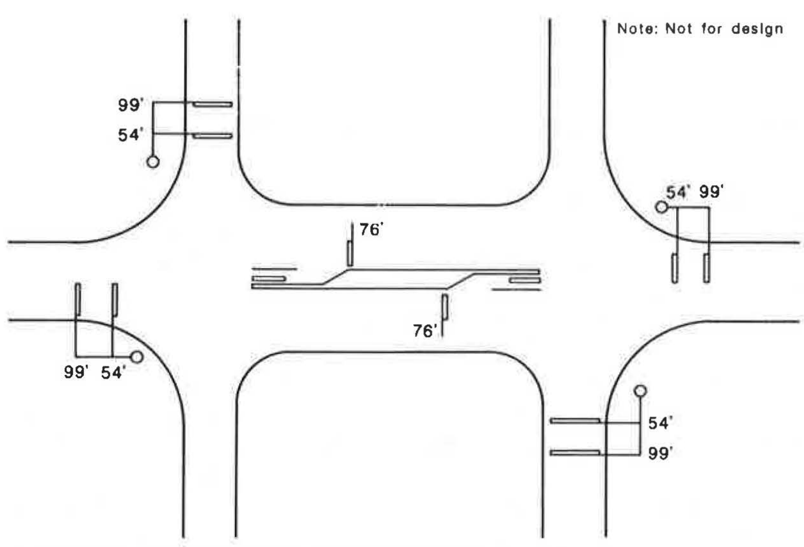


FIGURE 2 An illustration of three-phase detector layouts.

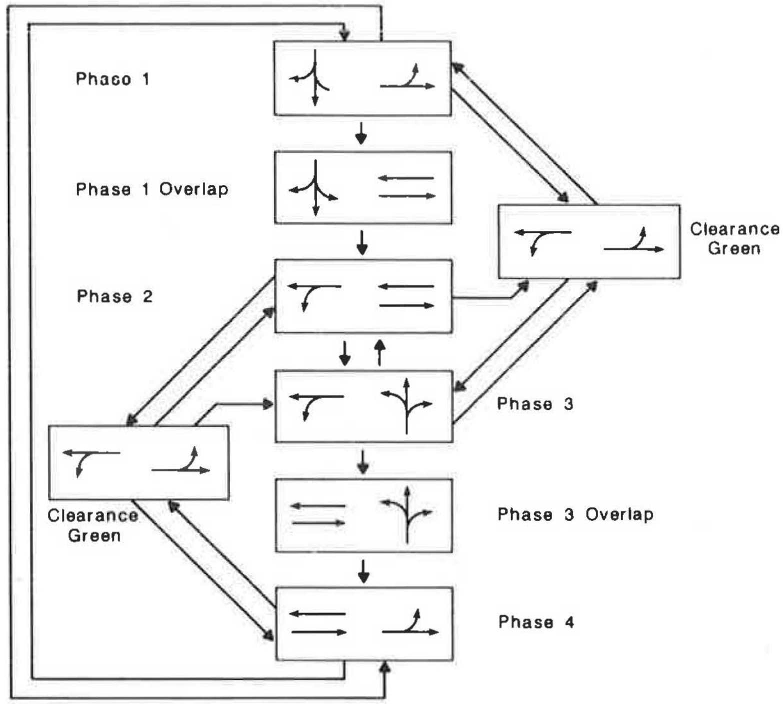


FIGURE 3 Four-phase, full-traffic-actuated diamond interchange phasing.

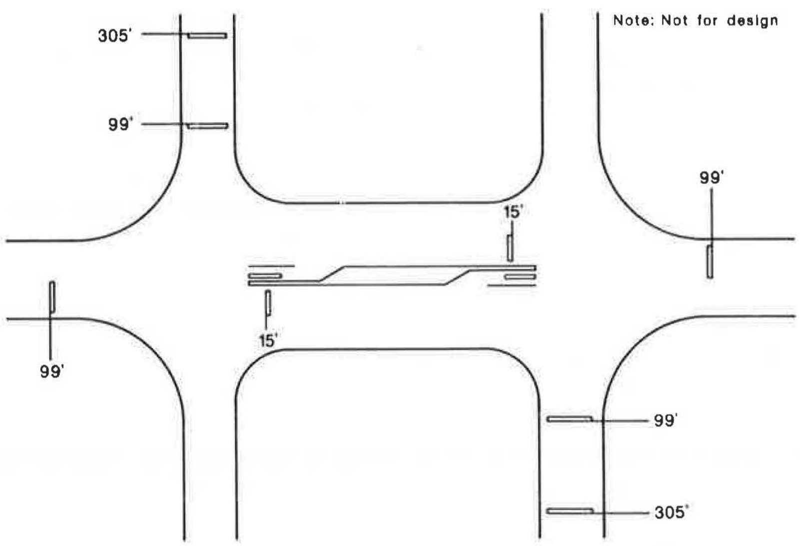


FIGURE 4 An illustration of four-phase, single-point detector layouts.

phase control. Other phase numbering schemes are also used in the literature.

Phase 1 in four-phase control is the lead, inbound frontage road phase. The choice of which frontage road leads is arbitrary. Phase 1 overlap is a fixed-duration phase equal to the travel time between intersections.

Proceeding clockwise around the interchange, Phase 2 is primarily an inbound, actuated, cross-street phase. Note, however, that only one arterial approach at a time initially receives the green.

Phase 3 likewise is the other frontage road movement. This phase operates similarly to Phase 1 and is followed by Phase 3 overlap.

Phase 4 concludes the services of actuated phases for this type of control. Phase 4 is the arterial inbound phase and is similar to Phase 2.

Detectors

Two detector configurations were also tested for four-phase control: (a) single-point and (b) multipoint detection. Figure 4 shows a typical detection plan for four-phase, single-point detection, whereas Figure 5 shows a common detector layout for four-phase with multipoint detection. Some variation in the detection plan was made at each site to best accommodate each interchange's geometrics and approach speeds.

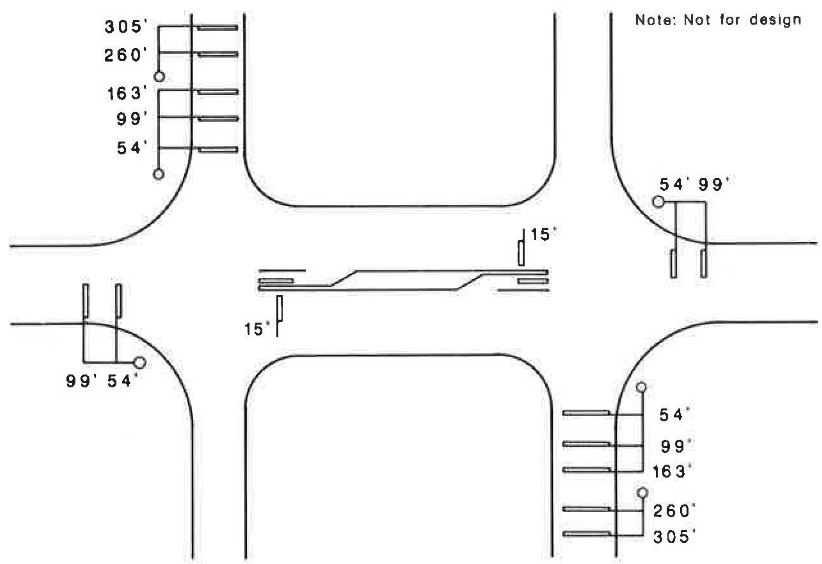


FIGURE 5 An illustration of four-phase, multipoint detector layouts.

Because of high volumes of low speed, turning traffic observed on the cross-street inbound approaches, practically no variation in single-point and multipoint detection configurations on the cross street was tested with three-phase or four-phase control at an interchange. In four-phase, single-point detection, one detector (set) was used at about 100 ft from the stop bar, the same as three-phase. In four-phase, multipoint detection, an additional detector was placed about 50 ft from the stop bar, which provided better signal change protection and shorter minimum greens but more actuations and only a slight reduction in gap timing for promoting gap-out.

Multipoint detection on the frontage roads used the five-detector system shown in Figure 5. The three detectors located closer to the intersection are connected to one amplifier. This special detector amplifier's output is routed through an external logic card to process inputs. When speeds of 40 mph (or related occupancy time) are recognized, this detector set is disabled by the logic card, and phase extension immediately switches to the upstream extension set of detectors. The upstream detectors will extend the green when headways of 2.1 to 2.5 sec are maintained and provide protection against possible dilemma zone problems for speeds up to 55 mph. Using these upstream detectors, gap-out for Phase 1 termination usually occurs at the desired time such that the end of the platoon arrives at the stop bar at the termination of Phase 1 overlap. The detector switching thus effectively promotes full utilization of Phase 1 overlap.

EXPERIMENTAL PLAN

Successful field studies were conducted at three diamond interchanges in Texas. Both three-phase and four-phase control systems were tested at each site. A control system is considered to be one type of controller phasing combined with one type of detector plan. The scope of the study limited field observations to only 1 day per type of signal control system studied per interchange.

Interchange Characteristics

The three sites offered a typical variety of geometric and traffic patterns for Texas. Some geometric commonality was also present. All interchanges provided continuous, one-way frontage road operations in a suburban environment. All interchanges were traffic actuated with each city having some experience with three-phase and four-phase control. All three interchanges could provide three-phase and four-phase control with existing equipment. However, the four-phase control tested used a special NEMA four-phase controller that had to be temporarily installed to provide single-point and multipoint detection. Three-phase control used the existing controller units.

A summary of selected diamond interchange attributes is given in Table 1. The schematic layout of Ocean and Pine

TABLE 1 INTERCHANGE SITE CHARACTERISTICS

Interchange Cross-street File Name	Dimensions Curb-to-Curb		Queue Storage Distance	Turnaround Lanes Present	Left-Turn Lane
	Outside	Inside			
North	232	160	150	No	No
Ocean	382	310	290	Yes	Yes
Pine	470	396	360	Yes	Yes

Streets were similar. Both had turnaround lanes on both sides. However, North Avenue was a fairly small interchange, Ocean Avenue was moderately sized, and Pine was a large interchange. Interchange lengths (distance along the cross street) ranged from 278 to 470 ft. North Avenue was the only one studied that was on a cross-street bridge at grade with the frontage roads. No left-turn or U-turn lanes were provided on the bridge.

The data collected at each site contained three types of measures: (a) traffic demand variables, (b) interchange control and geometric attributes, and (c) traffic performance measures. Field observations of system activity together with incidental records were also maintained in a log book for each day of the study.

Statistical considerations of randomness, stability, and sample size combined with previous experiences led to the selection of a 15-min time interval as being the time base for study. Each 15-min period was considered one independent study, or data point. Volumes and system performance (delay) queue counts to be described in following sections were obtained for each 15-min interval.

Traffic Volume

Traffic volume was used as the primary input variable. Traffic counts were made at each intersection by turning movement using manual observers. Time-lapse turning movement recorders with assistant recorders attached were used to initially record the turning movement counts by approach lane. Turning movement summaries were prepared for each approach by lane. The maximum volume [expressed in vehicles per hour per lane (vphpl)] observed on each approach for each 15-min study period was identified. These six "critical" volumes, three at each intersection, were then added together to form an "interchange total critical volume" for each study interval. That is

$$V = V_{1c} + V_{2c} + V_{3c} + V_{4c} + V_{5c} + V_{6c} \tag{1}$$

where V is the interchange total critical lane volume, vphpl, and V_{ic} is the critical lane volume on approach i, vphpl. Subscripts 1, 2, and 3 relate to the three approach legs on one intersection; whereas Subscripts 4, 5, and 6 relate to the corresponding movements on the other intersection. V_{3c} and V_{6c} represent the larger of the outbound through or left-turn flows at the respective intersections.

Computer programs were prepared during the data reduction phase to automatically make these critical volume determinations and summarize the total interchange results.

Cycle Length

Cycle length was measured for each study period and tested as a dependent variable and as an independent variable at various stages of the analysis process. Cycle length for actual control changes with each succeeding phasing sequence. Unlike pretimed control, the time of each cycle length for basic actuated control depends on short-term traffic volumes, number of phases, and traffic controller settings of (a) initial green, (b) gap extension, and (c) maximum green for each phase, together with other factors. An average cycle length over each 15-min period was determined by averaging the cycle lengths recorded by an observer.

Queue Delay

Signal efficiency is normally described in terms of delay, delay per vehicle or, as in the *1985 Highway Capacity Manual (5)*, in terms of stopped delay per vehicle. Stopped delay per vehicle on an approach serving an arrival flow of "v" vehicles per hour is

$$d = \frac{q}{v} \tag{2}$$

where

d = stopped delay, sec/veh;
q = average number of vehicles stopped in queue at an approach to the interchange during the study interval, veh, and
v = approach flow, veh/sec.

For each approach to the interchange, counts of the number of vehicles stopped in each lane for each approach were recorded every 15 sec. Averages by lane per approach were then determined for each of the 60 (4 × 15) samples over the respective 15-min period. A maximum average queue per lane per approach was then obtained. Maximum queues per lane per approach were determined during data reduction. Each maximum (critical) queue per lane per approach was denoted c_i.

Total interchange critical queue was used as the traffic control system performance measure of operational efficiency. Total interchange queue was derived from the six approaches similar to total input volume. Total interchange critical queue for a 15-min period is equal to

$$Q = Q_{c1} + Q_{c2} + Q_{c3} + Q_{c4} + Q_{c5} + Q_{c6} \tag{3}$$

where Q is the total interchange critical queue, veh/lane, and Q_{ci} is the maximum queue per lane on approach i, veh/lane.

Comparisons between system design attributes can be effectively made at the same total volume levels. However, Equation 2 indicates that comparisons of observed queues for different control systems cannot be made at different volume levels because the case having higher total interchange queue could have higher volumes, but less average delay per vehicle.

Study Plan

Field studies at the three interchanges were conducted from Tuesday through Friday during the Spring and Summer of 1984. A typical field study team was composed of eight field observers plus one study supervisor. Three study periods per day were provided to sample a wide range of volume levels and traffic patterns. A typical daily schedule ran from 7:30 a.m. to 9:00 a.m. followed by a breakfast break. A 2-hr study of off-peak and noon-hour traffic began at 11:00 a.m. and lasted until 1:00 p.m. Several traffic patterns occur during this period. Following lunch and a brief break, the afternoon study lasted from 4:30 p.m. until 6:00 p.m. Again, 15-min study intervals were obtained by all staff synchronizing their watches before each study. The study plan thus provided 5 hr of observation time per day with four data points per hour for a total of 20 (5 × 4 = 20) data points obtained per system configuration per interchange.

Data Reduction

Three levels of data reduction were performed. All manually recorded queue and turning movement counts were routinely logged following each study period. Dates and station locations were checked for accuracy. All turning movement counts were transferred from the counter boards to data sheets before departure from the site.

A considerable quantity of data had to be manually reduced in the office by staff personnel. Queue counts, in particular, required substantial time. Queue counts were being recorded on scribble pads at six approaches by lane every 15 sec. This sampling rate results in about 1,000 queue samples for all lanes during each 15-min study period, or a total of about 86,000 samples per interchange. All of these data points had to be manually tallied, averaged, and tabulated for coding into the computer.

The study data were coded into the Amdahl computing system at Texas A&M University using remote job entry WYLBUR terminals. Routine statistical summaries were prepared for each data set for visual inspection of the data for any apparent coding errors. Range and limit tests were conducted to further check for coding errors. Preliminary testing revealed that each data set contained consistent and expected trends in attributes. The data were then pooled to evaluate the performance characteristics of the four alternative diamond interchange control systems.

Data Analysis

The pooled data were analyzed by using statistical analysis techniques. The Statistical Analysis System *(6)* was used throughout the data analysis phase. Basic summary and descriptive statistics were used to illustrate diamond interchange traffic and queue characteristics. Further, multiple regression models and general linear hypothesis testing were used to evaluate the different signal phasings and detector configurations at the diamond interchanges. The detailed analysis techniques used, variables considered, and the evaluation processes followed to select the models describing the diamond interchange operational characteristics will be presented later.

STUDY RESULTS

The derived performance characteristics of four alternative diamond interchange control systems introduce the study results. These performance characteristics will be represented by a series of models or graphs illustrating relationships such as cycle versus critical volume, and critical queue versus critical volume and traffic pattern. Subsequently, alternative control systems, given either phasing plan or detection scheme, will be presented to illustrate performance differences. In the following sections, the four signal control systems are denoted as follows:

3S = three-phase, single-point detection
3M = three-phase, multipoint detection
4S = four-phase, single-point detection
4M = four-phase, multipoint detection

Cycle Length Versus Critical Volume and Traffic Pattern

The four alternative control systems were evaluated to determine the average cycle length that would be expected given the critical volume at the diamond interchange. The models developed are as follows:

3S, $C = 21.8 + 14.4\ (V/1{,}000)$, $R^2 = 0.68$
3M, $C = 20.8 + 13.5\ (V/1{,}000)$, $R^2 = 0.64$
4S, $C = 27.7 + 31.7\ (V/1{,}000)$, $R^2 = 0.72$
4M, $C = 21.5 + 25.4\ (V/1{,}000)$, $R^2 = 0.73$

where C is the cycle length in seconds and V is the sum of critical lane volumes at the interchange, vphpl.

Operating cycle lengths were found to increase with critical volume, as expected. The effect of traffic pattern was studied in the next step of the analysis process.

Because not only traffic volume but also traffic pattern affect cycle length, other variables representing traffic pattern were individually added to the model. The best models found from stepwise regression are as follows:

3S, $C = 14.5 + 14.4\ (V/1{,}000) + 20.1\ \text{RILCVE}$, $R^2 = 0.80$
3M, $C = 15.9 + 12.9\ (V/1{,}000) + 15.8\ \text{RILCVE}$, $R^2 = 0.76$
4S, $C = 38.7 + 32.3\ (V/1{,}000) + 33.8\ \text{RILCVE}$, $R^2 = 0.79$
4M, $C = 16.9 + 25.8\ (V/1{,}000) + 11.4\ \text{RILCVE}$, $R^2 = 0.75$

where RILCVE = internal left-turn volumes per sum of external critical volumes. Figure 6 shows the relationships found between cycle length and critical volume for the range of volumes studied using RILCVE = 0.4, the mean of the field studies. Several observations determined from Figure 6 are as follows:

1. Three-phase, multipoint detection consistently produced the shortest cycle length given traffic conditions.
2. Three-phase, multipoint detection had little advantage in cycle length when compared to three-phase, single-point detection.
3. Four-phase, single-point detection generated the longest cycle length given traffic conditions.
4. Four-phase, multipoint detection provided substantial reduction in cycle length as compared with four-phase single-point.
5. Three-phase control produced shorter cycles than did four-phase control.

Critical Queue Versus Critical Volume and Traffic Pattern

The effect of critical volume on critical queue for alternative control schemes was evaluated. The models developed are as follows:

3S, $Q = 1.12 + \text{Exp}\ [0.87\ (V/1{,}000)]$, $R^2 = 0.79$
3M, $Q = 1.22 + \text{Exp}\ [0.85\ (V/1{,}000)]$, $R^2 = 0.74$
4S, $Q = 1.75 + \text{Exp}\ [0.88\ (V/1{,}000)]$, $R^2 = 0.74$
4M, $Q = 1.09 + \text{Exp}\ [1.06\ (V/1{,}000)]$, $R^2 = 0.79$

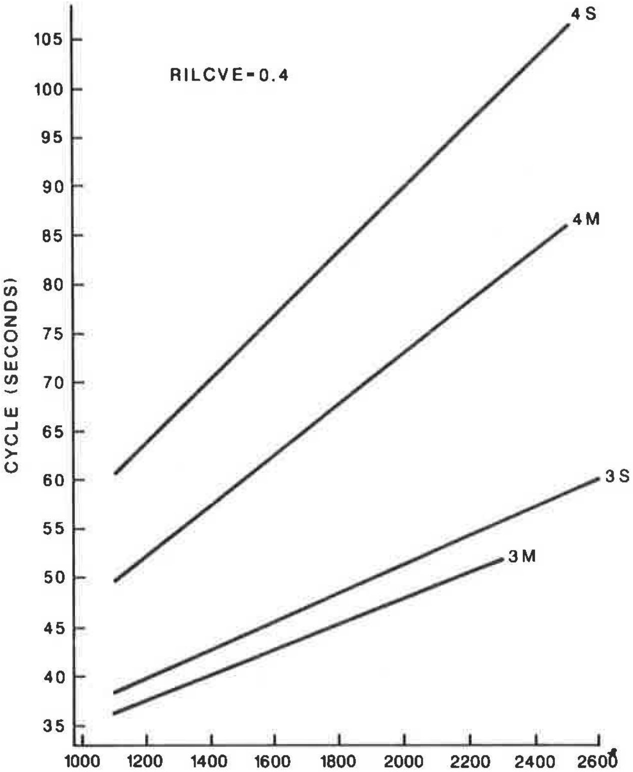


FIGURE 6 Relationship between cycle length and critical volume.

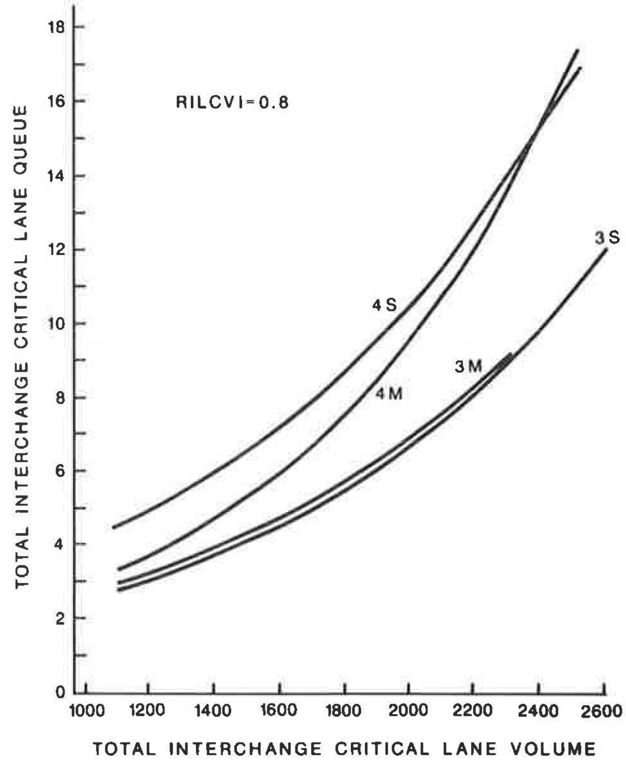


FIGURE 7 Relationships between critical queue versus critical volume and traffic pattern for RILCVI of 0.8.

As expected, these models predict that average queue increases with increasing critical volume.

The effect of different traffic pattern, in addition to traffic volume, on interchange queue performance was evaluated. Several variables previously explained to define traffic pattern at an interchange were individually tested. The best models found are as follows:

3S, $Q = 0.72 + \text{Exp}\ [0.976\ (V/1{,}000) + 0.35\ \text{RILCVI}]$, $R^2 = 0.84$

3M, $Q = 0.70 + \text{Exp}\ [0.938\ (V/1{,}000) + 0.50\ \text{RILCVI}]$, $R^2 = 0.89$

4S, $Q = 1.39 + \text{Exp}\ [0.943\ (V/1{,}000) + 0.17\ \text{RILCVI}]$, $R^2 = 0.76$

4M, $Q = 0.67 + \text{Exp}\ [1.185\ (V/1{,}000) + 0.36\ \text{RILCVI}]$, $R^2 = 0.84$

where RILCVI = internal left turns per sum of critical internal lane volumes.

Figure 7 shows the effect of traffic volume and traffic pattern at an interchange on traffic delay experienced for the range of volumes studied, using the mean RILCVI = 0.8 observed in the field studies. Several observations can be derived from Figure 7 as follows:

1. There was no significant difference in queue performance between three-phase, single-point, and multipoint detection given traffic volume at an interchange.
2. The four-phase, single-point system generated the highest delay among other alternative control schemes for a given traffic volume.
3. Three-phase control produced less delay than four-phase control given traffic volume at an interchange.

Three-Phase Control Detector Configurations

Figure 8 shows the queue performance characteristics between single-point and multipoint detection for three-phase control. Traffic pattern, given in terms of RILCVI, is shown at 0.4 and 1.0. It can be observed from Figure 8 that there appears to be no significant difference between single-point and multipoint detection for three-phase control at a given traffic volume and traffic pattern.

A general linear hypothesis test was performed to evaluate whether queue performance between single-point and multipoint detection for three-phase control was statistically different. No significant difference in queue performance was detected between single-point and multipoint detection for three-phase control.

Four-Phase Control Detector Configurations

Figure 9 shows the queue performance characteristics derived for single-point and multipoint detection for four-phase control. Traffic pattern is also depicted at 0.4 and 1.0 values of RILCVI. It can be observed from Figure 9 that multipoint detection for four-phase control generated shorter delay except when heavy

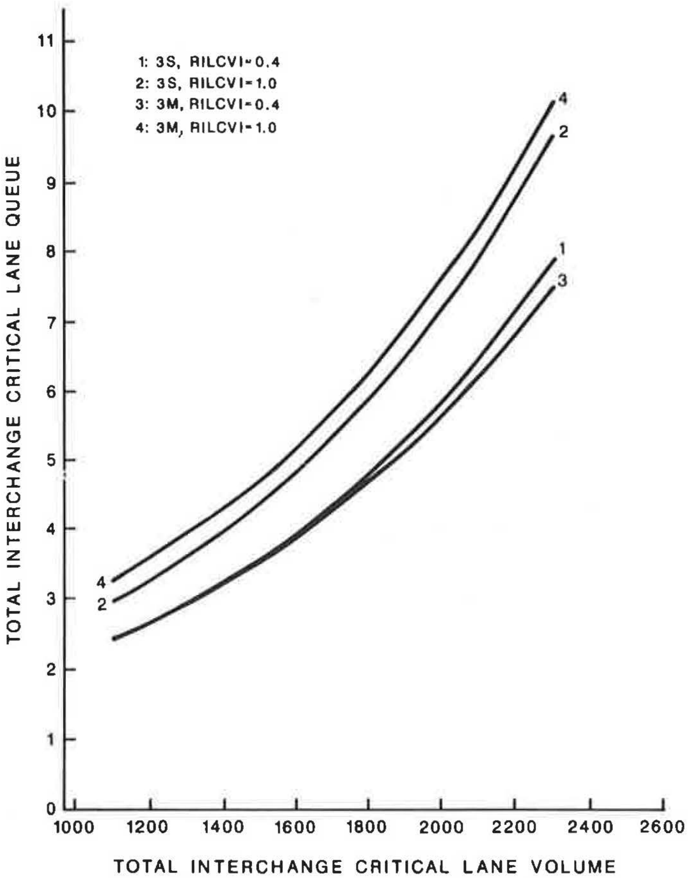


FIGURE 8 Queue performance characteristics between single and multipoint detection for three-phase control.

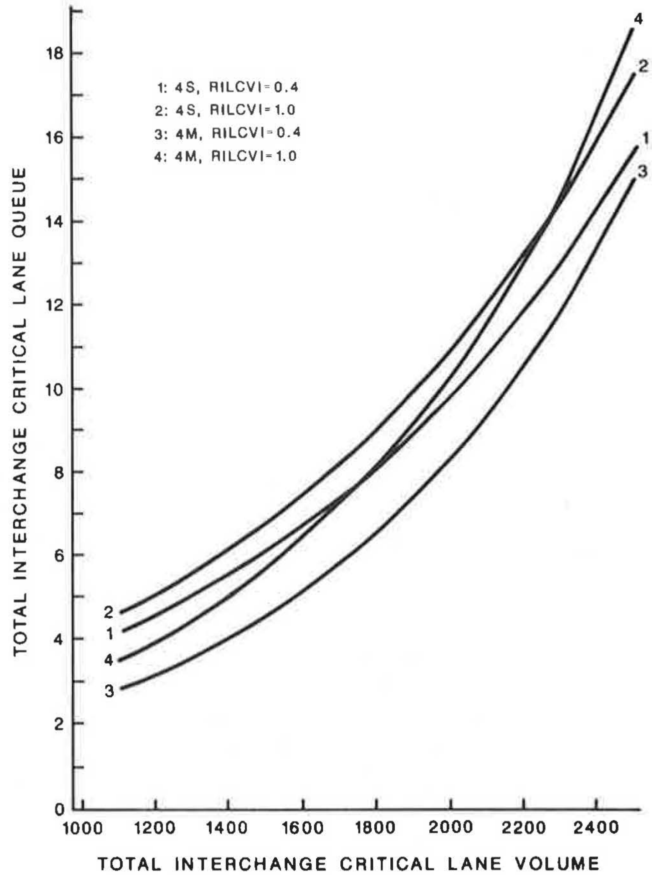


FIGURE 9 Queue performance characteristics between single and multipoint detection for four-phase control.

traffic flow together with heavy internal left turns exist at an interchange.

CONCLUSIONS AND RECOMMENDATIONS

The operational performance of traffic-actuated, signalized diamond interchange control systems has been examined in this study. Basic traffic-actuated controller units were used. All interchanges were operated isolated from all other intersections or interchanges. None of the interchanges was located within frontage road progressive systems. A wide range of volume levels were observed, but excessively heavy (or over capacity) volumes were infrequently observed, if at all.

Conclusions

The following conclusions were drawn from the data collected and field observations made in this study. They apply within the volume levels measured, traffic patterns experienced, and operational environment of one-way frontage roads in an urban area using basic actuated signal control.

1. Single-point detection is the more cost-effective three-phase detection system because it provides the same effectiveness as does the more costly multipoint detection system.

2. Multipoint detection is the more delay-effective, four-phase detection system. It provides more effectiveness but with a more expensive system. Its true cost-effectiveness is unknown.

3. Shorter cycle lengths are, in general, a desirable attribute for isolated interchange control. Phase terminations should be "snappy," with prompt phase termination becoming more critical as volume increases.

4. Four-phase control characteristically operates at a longer cycle length than does three-phase for a given traffic volume, but provides superior internal progression within the interchange.

5. Three-phase control can produce less overall queuing delay than four-phase for the same volume and level of detection. In most cases, however, this lower delay arises at a price of undesirable secondary stops within the interchange.

6. Three-phase control can be a good phasing strategy under selective geometric, traffic and control conditions. Three-phase works better when the interchange is wide and there is a high proportion of through flow, either on the frontage roads or on the cross street, or on both. In most cases, three-phase requires the use of relatively short cycle times with wider interchanges permitting better phase flexibility and smoother flow through the interchange.

7. Four-phase is an acceptable signal phasing strategy for typical urban interchange applications. Control stability and progressive flow are routinely provided but usually at a price of

increased cycle length and overall interchange delay unless the control is finely tuned.

8. Single-point detection produces, in general, longer cycle lengths than does multipoint detection. The trend toward longer cycle times for single-point detection is greater for four-phase than for three-phase control. Multipoint detection also can become susceptible to producing long cycle lengths under some heavy volume conditions.

Recommendations

The following recommendations are offered based on the results of this study. These recommendations apply to situations in which the signalized diamond interchange is operated isolated from all adjacent interchanges or intersections and the inside-to-inside, curb-to-curb dimensions between the frontage roads are 450 ft or less. In addition, only basic, full-actuated traffic signal controller units using small-area (point) detection are considered.

1. Single-point detection should be considered as a basic system component for three-phase control.
2. Multipoint detection on the frontage roads should be considered as a basic system component for four-phase control.
3. Four-phase with overlap control should be considered as a viable alternative in all cases of isolated, diamond interchange control where one-way frontage roads exist.
4. Three-phase control should be considered a viable alternative when any of the following isolated interchange control conditions exist:

a. When there is a small percentage of left-turn traffic on the frontage roads; or

b. When the interchange has sufficient internal queue storage capacity to store traffic without locking-up the turning movements within the interchange; or

c. When the interchange experiences freeway exit ramp or frontage road backup such that the backup affects freeway operation; and

d. The cycle length is kept short, phase termination snappy, and adequate visibility of the interchange signal operations exists.

5. Traffic control techniques should be considered for implementation at actuated diamond interchanges that delay phase calls and rapidly gap-out phases of lighter traffic in heavier traffic-demand situations. At high-volume interchanges, control features such as traffic-responsive, variable timings may be desired to reduce delays and minimize phase max-out even for multipoint detection.
6. There is a need to develop standard field test procedures for determining when an actuated diamond interchange controller unit is optimally fine-tuned to existing traffic conditions.
7. A traffic controller unit providing a combination of three-phase and four-phase operations could efficiently service a wide range of traffic and geometric conditions. The additional feature of providing improved progression along the cross street or frontage roads, or both, would be an additional attractive feature.

ACKNOWLEDGMENTS

This paper was prepared from research conducted within and supported by the Federal Highway Administration, U.S. Department of Transportation, and the Texas Department of Highways and Public Transportation. The authors wish to express their appreciation to the sponsoring agencies and to those who made major technical and administrative contributions to the study.

REFERENCES

1. C. Pinnell and D. C. Capelle. Operational Study of Signalized Diamond Interchanges. *HRB Bull. 324*, HRB, National Research Council, Washington, D.C., 1962, pp. 38–72.
2. P. K. Munjal. An Analysis of Diamond Interchange Signalization. In *Highway Research Record 349*, TRB, National Research Council, Washington, D.C., 1971, pp. 47–64.
3. C. J. Messer and D. J. Berry. Effects of Design Alternatives on Quality of Service at Signalized Diamond Interchanges. In *Transportation Research Record 538*, TRB, National Research Council, Washington, D.C., 1975, pp. 20–31.
4. *Diamond Interchange Program—Users Manual.* Report FHWA-IP-80-4. FHWA, U.S. Department of Transportation, Dec. 1980.
5. *Special Report 209: Highway Capacity Manual.* TRB, National Research Council, Washington, D.C., 1985.
6. *SAS User's Guide.* SAS Institute Inc., Cary, N.C., 1979.

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

Conversion from Permissive to Exclusive/ Permissive Left-Turn Phasing: A Before-and-After Evaluation

ANNE STONEX AND JONATHAN E. UPCHURCH

A before-and-after study was conducted to determine the effects of converting left-turn signal phasing from a permissive condition to an exclusive/permissive condition. Data collection was conducted in April 1984 and February 1985. Time-lapse photography was used to collect data on the numbers of vehicles already stopped, stopping, or not stopping at 5-sec intervals. Each movement (left-turn and through) and direction were separately recorded. These data, in turn, were analyzed to determine traffic volumes, average and total amounts of vehicle stopped delay, the percentage of vehicles stopping, and the percentage of left-turn vehicles. Mean values of these factors for before and after data were compared to determine the significance of any differences. The results showed that left-turn volumes increased significantly in the after phase. However, when these volumes were expressed as a percentage of total volume (which also increased), the increases were not significant. The percentage of vehicles that stopped increased dramatically from 43 percent of all vehicles in the before phase to 71 percent of all vehicles in the after phase. Average delays to southbound through traffic more than quadrupled in the after phase, whereas those to northbound through traffic more than tripled. Average delay to left-turn vehicles decreased to 82 percent of the before values; not a statistically significant amount. The conversion resulted in 87.9 veh-hr of additional delay per day. This delay converts to a cost of $398,587/year in additional vehicle operating, travel time, and vehicle emissions costs. Longer cycles, loss of progression, and inefficient use of green time increased the number of stopping vehicles and vehicle delay. The improvements in processing left-turn vehicles were obtained at the expense of inconveniencing the through movement. A comparison of before- and after-accident experience was not included in this study.

Described in this paper is a before-and-after study that evaluated the effects of converting left-turn signal phasing from a permissive condition to an exclusive/permissive condition.

Three types of left-turn phasing are in general use:

- *Permissive left turn.* Vehicles are allowed to make a turn on a circular green indication but must yield to opposing traffic.
- *Exclusive left turn.* Vehicles are allowed to make a turn only on a green arrow indication and have the right of way while the green arrow is displayed.
- *Exclusive/permissive.* Vehicles are allowed to make a turn either on a green arrow indication or on the circular green after the green arrow has been terminated and after yielding to oncoming traffic.

A. Stonex, 39499 Chart St., Mt. Clemens, Mich. 48045. J. E. Upchurch, Department of Civil Engineering, Arizona State University, Tempe, Ariz. 85287.

At present, there is no uniform method of application of left-turn phasing throughout the United States. A 1985 survey by the Colorado-Wyoming section of the Institute of Transportation Engineers drew the response of 218 jurisdictions. These jurisdictions listed 175 different criteria (based on delay, accidents, volumes, or other factors) for installing exclusive or exclusive/permissive phasing *(1)*. The multitude of different criteria being used strongly suggests that additional data on the effects of different types of left-turn phasing are needed to develop more uniform methods of application.

In addition to the Colorado-Wyoming study, other researchers have summarized current practice or have conducted studies to develop criteria or warrants. Agent and Deen prepared an excellent summary of state warrants or guidelines in 1978 *(2)*. Mohle and Rorabaugh documented the effects of installing exclusive left-turn phasing *(3)*. Warren reported on accident experience *(4)*. Upchurch used matched pairs of intersections to determine delay and other impacts *(5)*.

Few conversions from permissive to exclusive/permissive phasing are documented in the literature. Most of the intersections used as the subject of before-and-after studies were changed from permissive to exclusive or vice versa. Upchurch's work *(5)*, using matched pairs of intersections, produced much useful information. However, it was not possible in that study to duplicate intersection geometry, cycle length, turning movement percentages, and vehicle arrival patterns. These observations suggested that a before-and-after study of a permissive to exclusive/permissive conversion would be very useful. Before-and-after data collection at one location would minimize the number of confounding factors in the analysis.

The city of Phoenix has recently developed a strong interest in the subject of left-turn phasing. Political and engineering decisions in 1984 led to the opportunity to conduct a before-and-after study.

Phoenix has an excellent 1-mi grid system of major arterial streets, most of which are six or seven lanes wide. Arterials are heavily relied on because the city has fewer miles of freeway per capita than any other urban area of its size. Dramatic population and traffic growth have strained the surface street system. A frustrated public has come to believe that a left-turn arrow is the quick-and-easy solution to the problem. This belief is so popular and widespread that a mayoral candidate (subsequently elected) adopted "more left-turn signals" as part of his campaign agenda.

Phoenix's traffic engineering department has historically been reluctant to use exclusive or exclusive/permissive phasing because it (a) takes away green time from the through movement, (b) increases overall intersection delay and reduces capacity, and (c) disrupts the progression that works exceptionally well on the 1-mi grid system. Intuitively, the traffic engineers believed that left-turn arrows would have serious drawbacks. In response to political pressure, exclusive/permissive phasing was installed on an experimental basis at several locations to evaluate the impacts. One of these locations—44th Street and Thomas Road—was the site of this research project.

This intersection was chosen because it presented the opportunity to conduct a before-and-after study of the effects of the conversion from permissive to exclusive/permissive left-turn phasing. However, the researchers could not control or eliminate the other factors (cycle length, loss of progression, etc.) that confounded the analysis. The scope of the study did not allow any modifications to the intersections needed to determine the individual contribution of each factor to changes in "after" intersection operations.

"After" data were collected 7 months after installation of exclusive/permissive phasing. Therefore, drivers had 7 months to become aware of and adapt to the new signal phasing.

An analysis of before-and-after accident experience was not included in the scope of the study. The after condition described in this paper existed for less than 1 year; significant changes in traffic signal progression were made at the end of the after period. The short after period prevented any conclusive analysis of accident experience. Driver surveys were not within the scope of the study.

OBJECTIVE

The overall objective of the study was to evaluate the changes in intersection operation as a result of the installation of exclusive/permissive phasing at 44th Street and Thomas Road. Specific objectives were to answer the following questions:

1. What was the net change in delay to the 44th Street approaches, considering both through and left-turn vehicles?
2. Did delay to left-turn vehicles increase or decrease with the addition of exclusive/permissive phasing?
3. Did delay for nonturning vehicles increase?
4. What was the effect of the change in left-turn phasing on the ratio of green time to cycle length for the through movement?
5. What effect did left-turn arrows have on vehicle operation cost, air polluting emissions, fuel consumption, and person-hours of travel?
6. Have left-turn volumes increased as a result of the installation of exclusive/permissive phasing?

INTERSECTION DESCRIPTION

Forty-Fourth Street and Thomas Road are both major arterials with three through lanes in each direction and left-turn bays on each approach. The left-turn bays on the 44th Street approaches are about 200 ft long with storage space for 10 vehicles. Concrete medians channelize the traffic on all approaches. Signal heads are mounted on poles in the medians and on the corners as well as overhead. The signal heads in the medians were changed from the standard two-phase, three-section type to a five-section stacked type that includes green and yellow arrows.

Data were collected, via time-lapse film, only for the 44th Street approaches (northbound and southbound). Data were not collected on the Thomas Road approaches because of limited resources. In the before phase, the 44th Street average daily traffic (ADT) was 43,500 vehicles per day. The basic traffic patterns in the after phase were quite similar to those in the before phase. The before-and-after volumes for each direction and movement by hour of filming are given in Table 1.

Signal Timing

The signal timing for the before phase is given in Table 2. Cycle lengths ranged from 50 to 65 sec. North-south green time varied from 19 to 29 sec. The ratio of green time to cycle length (G/C ratio) ranged from 36.4 to 48.3 percent. The signals were two-phase, pretimed, and part of a progressive signal system. Progression speeds varied with cycle length from 28 to 36 mph.

The timing schedule for the after phase is given in Table 3. It should be noted that between 6:30 p.m. and 6:30 a.m., only permissive phasing is used on the north and south approaches. Except for this time period, G/C ratios for the through movement decreased in the after phase. This was because the increase in through green time was not as large as the increase in cycle length. G/C ratios for the through movement ranged from 36.4 to 48.3 in the before phase and from 28.5 to 36.7 in the after phase.

The cycle lengths shown are maximum values. Cycle lengths actually varied with demand throughout the day, but were not measured on a cycle-by-cycle basis. Upstream detectors determined the length of the through-clearance interval up to a maximum of 4.6 sec. The minimum arrow display time was 6 sec; the maximum, 10 sec.

Intersection Operation

In the after phase of the study, the intersection continued to operate with only permissive phasing from 6:30 p.m. to 6:30 a.m. From 6:30 a.m. to 6:30 p.m., the exclusive left-turn phase is actuated only if there are three or more vehicles in the left-turn lane. This is accomplished by using two detection loops in the left-turn bay. One loop is just behind the stop line and the second is 50 ft behind the stop line. The presence of vehicles on both detectors is required to call the exclusive phase. The exclusive phase is not called when there are less than three vehicles in the queue; in this case the through clearance interval can process two left-turn vehicles.

When used, the left-turn arrow leads the through-green phase. Overlaps are used if the exclusive phase is actuated for left turns in one direction but not the other.

The previously described operation of the exclusive phase was planned so as to limit the following disadvantages of the additional phase:

1. Loss of progression,
2. Increased cycle length,
3. Decreased G/C ratio for the through movement, and
4. Increased delay to through traffic.

Restricted hours of exclusive phase operation and minimum left-turn demand thresholds were efforts to limit its use. The main advantages of the addition of the exclusive phase were

1. Reduced left-turn delays,
2. Quicker dispersal of left-turn queues,
3. Additional left-turn capacity,
4. Reduced interference with through traffic due to "spillover" from the left-turn bay, and
5. Satisfaction of public demand to install left-turn arrows.

A comparison of traffic flows from the before-and-after films shows that the after-phase operation was neither as smooth nor as efficient as it was in the before phase.

A before-and-after type study was used to reduce, as much

TABLE 1 VOLUMES BY HOUR OF FILMING

	Northbound			Southbound		
Hour[a]	Before	After	Percent Difference	Before	After	Percent Difference
Through Volumes (vehicles per hour)						
1	942	949	+0.7	1,294	1,209	-6.6
2	904	1,036	+14.6	1,012	1,123	+11.0
3	1,037	1,221	+17.7	1,087	1,194	+9.8
4	1,309	1,267	-3.2	1,239	1,215	-1.9
5	1,243	1,367	+10.0	1,070	1,233	+15.2
6	1,330	1,513	+13.8	1,272	1,301	+2.3
7	1,498	1,790	+19.5	1,335	1,239	-7.2
8	1,748	1,798	+2.9	1,185	1,295	+9.3
Left-Turn Volumes (vehicles per hour)						
1	159	169	+6.3	131	140	+6.9
2	182	208	+14.3	129	156	+20.9
3	217	251	+15.7	154	172	+11.7
4	209	260	+24.4	160	203	+26.9
5	196	273	+39.3	158	176	+11.4
6	192	238	+24.0	151	168	+11.3
7	190	264	+38.9	143	185	+29.4
8	181	217	+19.9	139	195	+40.3

[a] Each hour represents one reel of exposed film. Filming began at about 8:15 a.m. and concluded at about 5:15 p.m.

TABLE 2 SIGNAL TIMING AND G/C RATIO BEFORE

Time of Operation	Cycle Length (sec)	N–S Green (sec)	N–S Yellow (sec)	Progression Speed (mph)	Green Time G/C ratio (%)
6:45 a.m.–8:15 a.m.	60	29	4	30	48.3
8:15 a.m.–4:00 p.m.	50	19	4	36	38.0
4:00 p.m.–5:10 p.m.	55	20	4	33	36.4
5:10 p.m.–5:40 p.m.	65	29	4	28	44.6
5:40 p.m.–6:00 p.m.	55	20	4	33	36.4
6:00 p.m.–6:45 a.m.	50	19	4	36	38.0

TABLE 3 SIGNAL TIMING AND G/C RATIO AFTER

Time of Operation	Cycle Length (sec)	N–S Through Green (sec)	N–S Through Yellow (sec)	N–S Left-Turn Arrow (sec)	N–S Left-Turn Yellow (sec)	Green Time for N–S Through (G/C ratio, %)
6:30 a.m.–9:00 a.m.	105.2 max	30	4.6	10 max	3	28.5
9:00 a.m.–3:30 p.m.	77.2 max	25	4.6	6 max	3	32.4
3:30 p.m.–6:00 p.m.	105.2 max	30	4.6	10 max	3	28.5
6:00 p.m.–6:30 p.m.	77.2 max	25	4.6	6 max	3	32.4
6:30 p.m.–6:30 a.m.	68.2 max	25	4.6	0	0	36.7

as possible, differences in intersection characteristics and operation when comparing permissive phasing to exclusive/permissive phasing. It is emphasized that some factors, which were beyond the control of the researchers, did change. The cycle length was increased, the ratio of through green time to cycle length was decreased, volumes increased slightly, and the previous pattern of traffic progression was disrupted. Although it would have been desirable for a perfect before-and-after comparison to have the same cycle length, progression patterns, volume, and ratio of north-south to east-west green time, these characteristics could not be controlled.

DATA COLLECTION

Time-lapse photography was used for data collection; it is the only practical data collection method for accurately obtaining information on volume and associated vehicle delay. The films were used to determine left-turn volumes, opposing volumes, and delay (both to left-turn and through vehicles).

The time-lapse camera was located on a lift truck adjacent to the right lane of the south approach and approximately 300 ft from the intersection. The camera was approximately 30 ft above the roadway. From this location the through and left-turn movements on the north and south approaches were observed and recorded on film.

Eight hours of film were exposed in both the before and after phases. The 8 hr covered a time period from about 8:15 a.m. to 5:15 p.m. (including short breaks for changing film). The traffic observed in the 8-hr period accounted for about 51 percent of ADT. A speed of one frame per second was used for all filming. Filming was continuous in order to be able to calculate delays based on 1-sec intervals. Each roll of film had 3,600 frames (50 ft roll) and ran for 1 hr. Filming was done for the before phase on Friday, April 13, 1984, and for the after phase on Friday, February 22, 1985.

DATA REDUCTION

The basic types of information obtained from the time-lapse films were volume and delay data. Stopped time delay was the specific type of delay calculated in this study. It measures the time a vehicle is stopped and does not include time losses caused by deceleration and acceleration. Wherever the term delay is used in this paper, it refers to stopped time delay.

The time-lapse film was projected using a time-lapse projector at a slow rate of speed. Viewing of the films, observation of vehicle movements, and tabulation of data resulted in the collection of data on volume, the number of vehicles stopping, the number of vehicles not stopping, total delay, average delay per stopped vehicle, average delay per approach vehicle, and the percent of vehicles that stopped. These data were collected separately for left-turn and through movements and for the near- and far-side approaches to the intersection. These data were tabulated for 5-min intervals.

Although the time-lapse film was exposed at a rate of one frame per second, 5-sec intervals were used for recording volume and delay data. This interval facilitated data reduction and analysis. A 5-sec interval of film was projected, and the number of vehicles that (a) were stopped, (b) came to a stop in that interval, and (c) did not stop at all while traversing the intersection were observed and tallied. A stopped vehicle was defined as one that was stopped and waiting for the signal to turn green or for a suitable gap (in the case of left-turn vehicles).

Stopped time delay was used for calculating delay. In this study, stopped vehicles were counted in 5-sec intervals. Every 5 sec, the number of vehicles stopped (in through or left-turn lanes) was recorded. The total delay (for all vehicles on the approach) was calculated as the total number of vehicles observed multiplied by the observation interval (5 sec).

Volume and delay data were summed for 5-min periods. Average delay per stopped vehicle, average delay per approach vehicle (vehicles on the approach), and the percent of vehicles that stopped were calculated from the volume and delay data. In addition, the percentage of all vehicles on an approach that turned left was calculated. Data were further summarized by summing the preceding factors for 1-hr periods.

STUDY FINDINGS AND ANALYSIS

Volume

Traffic volume increased on all four movements between April 1984 and February 1985. The number of vehicles per hour observed for each movement are listed in order of hour of filming in Table 1. This hourly breakdown shows the variations throughout the day. The 8-hr totals and overall average volumes are given in Tables 4 and 5.

The largest increases were in left-turn volume. Average hourly southbound left-turn volume increased from 146 to 174 vehicles, a 19.2 percent increase. The increase was statistically significant at the 95 percent level of confidence.

The increase in northbound left-turn volume is even greater.

TABLE 4 VOLUME—8-HR TOTAL IN VEHICLES

	Before	After	Percent Difference
Through volume (NB and SB)	19,505	20,750	+6.4
Left-turn volume (NB and SB)	2,691	3,275	+21.2
Southbound volume (through and left turn)	10,659	11,204	+5.1
Northbound volume (through and left turn)	11,537	12,821	+11.1
Total NB and SB volume	22,196	24,025	+8.2

NOTE: NB = northbound; SB = southbound.

TABLE 5 VOLUME—8-HR AVERAGE IN VEHICLES PER HOUR

	Before	After	Percent Difference
Southbound through	1,187	1,226	+3.3
Northbound through	1,251	1,368	+9.3
Southbound left turn	146	174	+19.8
Northbound left turn	191	235	+23.2

Hourly average northbound left-turn volumes increased by 23.2 percent from 191 to 235 vehicles per hour. The increase was statistically significant at the 95 percent confidence level. Total left-turn volume for both directions increased by 21.2 percent.

Hourly average southbound through volume increased by 3.3 percent from 1,187 to 1,226 vehicles per hour. This increase is not statistically significant. Northbound through volume rose from 1,251 to 1,368 vehicles per hour. The increases in through volume are not statistically significant.

The total north-south volume rose 8.2 percent from 22,196 to 24,025 left-turn and through vehicles in 8 hr.

Left-Turn Volume as a Percentage of Total Volume

The combined increases in southbound through and left-turn volumes over an 8-hr period raised total southbound volume by 5.1 percent from 10,659 to 11,204 vehicles (see Tables 4 and 6). Southbound left-turn volume grew from 10.9 percent to 12.5 percent of total (left-turn and through) southbound volume for a relative increase of 14.7 percent in the after phase. This increase is statistically significant at the 95 percent confidence level, but not at the 99 percent level.

TABLE 6 LEFT-TURN VOLUME AS A PERCENTAGE OF TOTAL VOLUME

	Before	After	Percent Difference
Southbound volume (through and left turn)	10,659 vehicles	11,204 vehicles	+5.1
Southbound volume (left turn only)	1,165 vehicles	1,395 vehicles	+19.7
Left-turn percentage of through + left	10.9%	12.5%	+14.7
Northbound volume (through and left turn)	11,537 vehicles	12,821 vehicles	+11.1
Northbound volume (left turn only)	1,526 vehicles	1,880 vehicles	+23.2
Left-turn percentage of through + left	13.2%	14.7%	+11.4
SB and NB volume (through and left turn)	22,196 vehicles	24,025 vehicles	+8.2
SB and NB volume (left turns only)	2,691 vehicles	3,275 vehicles	+21.7
Left turn percentage of through + left	12.1%	13.6%	+12.4

Total northbound volume rose by 11.1 percent, from 11,537 to 12,821 vehicles in 8 hr. The portion of this volume demanding left turns increased from 13.2 percent to 14.7 percent for a relative increase of 11.4 percent. The increase in northbound left-turn volume, expressed as a percent of total northbound volume, is not statistically significant.

Whether or not the relatively larger growth of left-turn demand is a result of the addition of left-turn phasing could only be answered by a survey of drivers. It is likely that the exclusive phase attracts some drivers because they perceive it as safer and more convenient. It may also attract drivers who previously used circuitous routes to avoid a lengthy left-turn delay.

Total Delay

Definite changes have occurred in the amount of total delay and its distribution. Table 7 gives 8-hr totals for delay in vehicle hours. The most dramatic changes have been the increases in through delay. Total southbound through delay for 8 hr increased from 11.59 to 49.34 vehicle-hours. The after value is 4.26 times the before value.

TABLE 7 DELAY—8-HR TOTAL IN VEHICLE-HOURS

	Before	After	Percent Difference
Southbound through	11.59	49.34	+326
Northbound through	14.72	52.31	+255
SB + NB through	26.31	101.64	+286
Southbound left turn	18.98	15.25	−19.6
Northbound left turn	42.96	35.26	−17.9
SB + NB through and left-turn combined	88.24	152.16	+72.4

Northbound through delay also increased markedly, from 14.72 vehicle-hours to 52.31 vehicle-hours. The after value is just over 3.5 times that of the before data.

The total through delay for the before phase is 26.31 vehicle-hours in 8 hr; for the after phase, 101.64 vehicle-hours or 3.86 times that of the before phase. Total volume in the after phase was only 1.08 times larger.

Left-turn delay decreased, as expected. Total southbound left-turn delay for 8 hr decreased by 19.6 percent, from 18.98 to 15.25 vehicle-hours. Total northbound left-turn delay decreased by 17.9 percent, from 42.96 to 35.26 vehicle-hours.

Total before north-south delay (through and left-turn movements combined) was 88.24 vehicle-hours in 8 hr. After the change in phasing and cycle length, total north-south delay was 152.16 vehicle-hours in 8 hr. The increase was 63.92 vehicle-hours or 72.4 percent.

The decrease in total left-turn delay did not offset the increase in total through delay. The through movements (86.3 percent of total north-south traffic) were penalized to benefit the left-turn movements that comprised only 13.7 percent of the total north-south volume.

The increase of 63.92 vehicle-hours of delay is for an 8-hr period during which filming was conducted. There are a total of 12 hr in the day during which exclusive phasing can be actuated. Based on relative volume levels in the 4 hr that were not filmed and the fact that volume increased 8.2 percent between the before and after phases, it is estimated that the total daily increase in delay as a result of the change in phasing is at least 87.9 vehicle-hours.

The use of green time for handling left-turn instead of through vehicles resulted in inefficient use of green time, which was a primary cause of increases in total delay. The longer cycle length caused the stopped through vehicles to be delayed longer than before; the loss of progression increased the number of vehicles forced to stop.

Average Delay

The impacts of the increase in total delay have already been discussed; however, the relationship of total delay to volume has not yet been explored. Average delays to through and left-turn movements for both directions are a direct expression of the volume/total delay interaction. Table 8 gives average delay, in units of vehicle-seconds per vehicle, for the 8 hr of filming.

Average delay to all southbound through vehicles increased from 4.4 to 18.1 vehicle-seconds per vehicle, a 311 percent increase. The corresponding change in northbound values was from 5.3 to 17.2 vehicle-seconds per vehicle, a 225 percent increase. Average delay to all through vehicles increased by 259 percent, from 4.9 to 17.6 vehicle-seconds per vehicle. All of these changes were shown to be statistically significant.

Average delay to southbound left-turn vehicles decreased from 58.6 to 39.4 vehicle-seconds per vehicle, a drop of 32.8 percent. The 32.8 percent decrease in average delay per vehicle is considerably more than the 19.6 percent decrease in total delay reported in the preceding section. This emphasizes the importance of examining average delay. Left-turn volume increased while total delay decreased, resulting in a much larger drop in average delay.

TABLE 8 DELAY—AVERAGES OVER 8-HR IN VEHICLE-SECONDS PER VEHICLE

	Before	After	Percent Difference
Southbound through	4.4	18.1	+311.4
Northbound through	5.3	17.2	+224.5
SB + NB through	4.9	17.6	+259.2
Southbound left turn	58.6	39.4	−32.8
Northbound left turn	101.3	67.5	−33.4
SB + NB left turn	82.8	55.7	−32.7

Average northbound left-turn delay was reduced from 101.3 to 67.5 vehicle-seconds per vehicle, a decrease of 33.4 percent. Again, the decrease in average delay is magnified by a concurrent rise in volume. The average delay to all left-turn vehicles (both directions combined) decreased from 82.8 to 55.7 vehicle-seconds per vehicle. Surprisingly, analysis of variance showed that all of the reductions in average left-turn delay were not statistically significant.

Graphical Analysis of Left-Turn Delay

Opposing Volume Ranges

Figure 1 shows average left-turn delay plotted as a function of opposing volume. One curve shows left-turn delay for the before phase; the other curve shows it for the after phase. The graph was constructed by partitioning average left-turn delays for 5-min intervals into ranges of opposing volume. The 5-min average delay values were used to calculate a mean left-turn delay for each volume range.

The before plot indicates a general tendency for average left-turn delays to increase with increasing opposing volume. Larger opposing volumes result in fewer gaps for left turns; thus, left-turn delay increases. The after plot shows a much narrower range of left-turn delay. The shape of the curve reflects the fact that larger opposing volumes cause longer left-turn queues. When such volumes persist, the exclusive phase may be called for each cycle. This hastens dispersal of the queues, thus reducing delay.

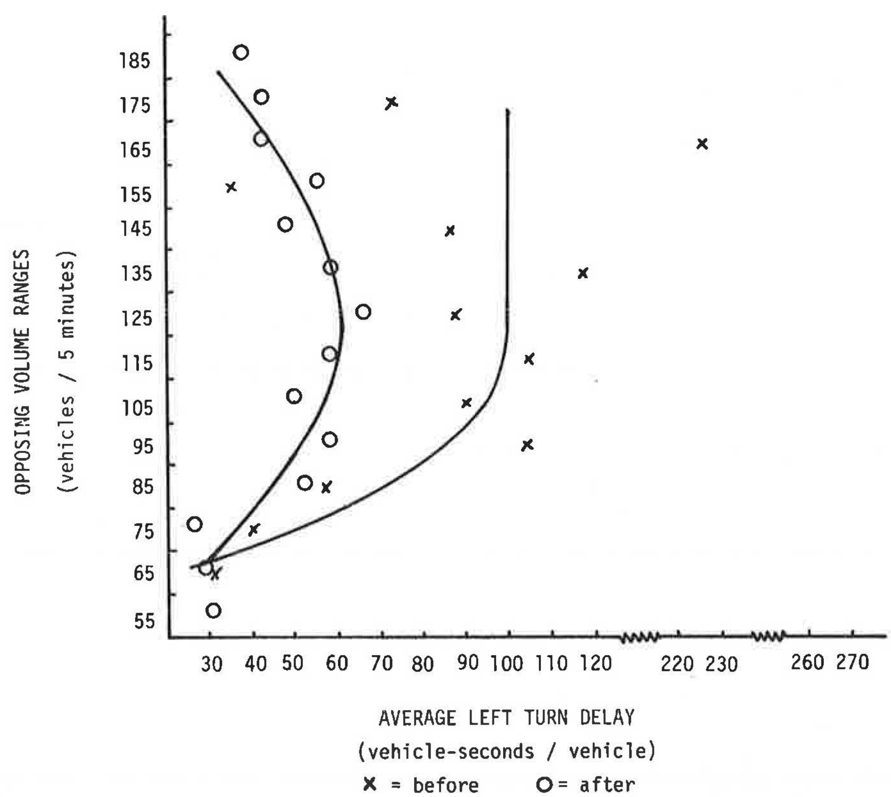


FIGURE 1 Plot of average left-turn delay versus ranging of opposing volume.

Volume Cross Product

As described, Figure 1 relates left-turn delay only to opposing volume. Figure 2 shows the relationship between average left-turn delay and volume cross product (VCP). Volume cross product is the opposing volume multiplied by the left-turn volume. As such, it is a simplified index of conflicts between left turn and opposing traffic. Volume cross product is a useful measure for relating average left-turn approach delay to traffic stream conditions.

The VCP ranges in Figure 2 represent increments of 200 vehicles^2/5 min. The only difference between the generation of this graph and that for the volume ranges (Figure 1) is the partitioning of average delays.

The data in Figure 2 indicate little change in average left-turn delay at low-volume levels (volume cross product of 500 to 700 vehicles^2/5 min). At these volumes the exclusive/permissive system functions like a permissive system because left-turn demand is low. As volume increases (higher-volume cross products), however, left-turn demand is sufficient to call the exclusive phase. As a result there is a significant reduction in left-turn delay.

Percent of Vehicles Stopping

The addition of the exclusive left-turn phase was not the only change made on 44th Street. As the cycle length was increased, the ratio of through-green time to cycle length was decreased. As a result of the change in cycle length, progression on 44th Street could no longer be achieved. The percentage of arriving vehicles that were forced to stop at the intersection increased.

In the before phase, the average fraction of southbound through traffic that stopped was 34.7 percent over 8 hr. This percentage nearly doubled in the after phase, to 64.5 percent stopping (see Table 9). The average percent of northbound through traffic that stopped also rose—from 35.7 to 67.7 percent. The increases in the percent of through traffic that stopped

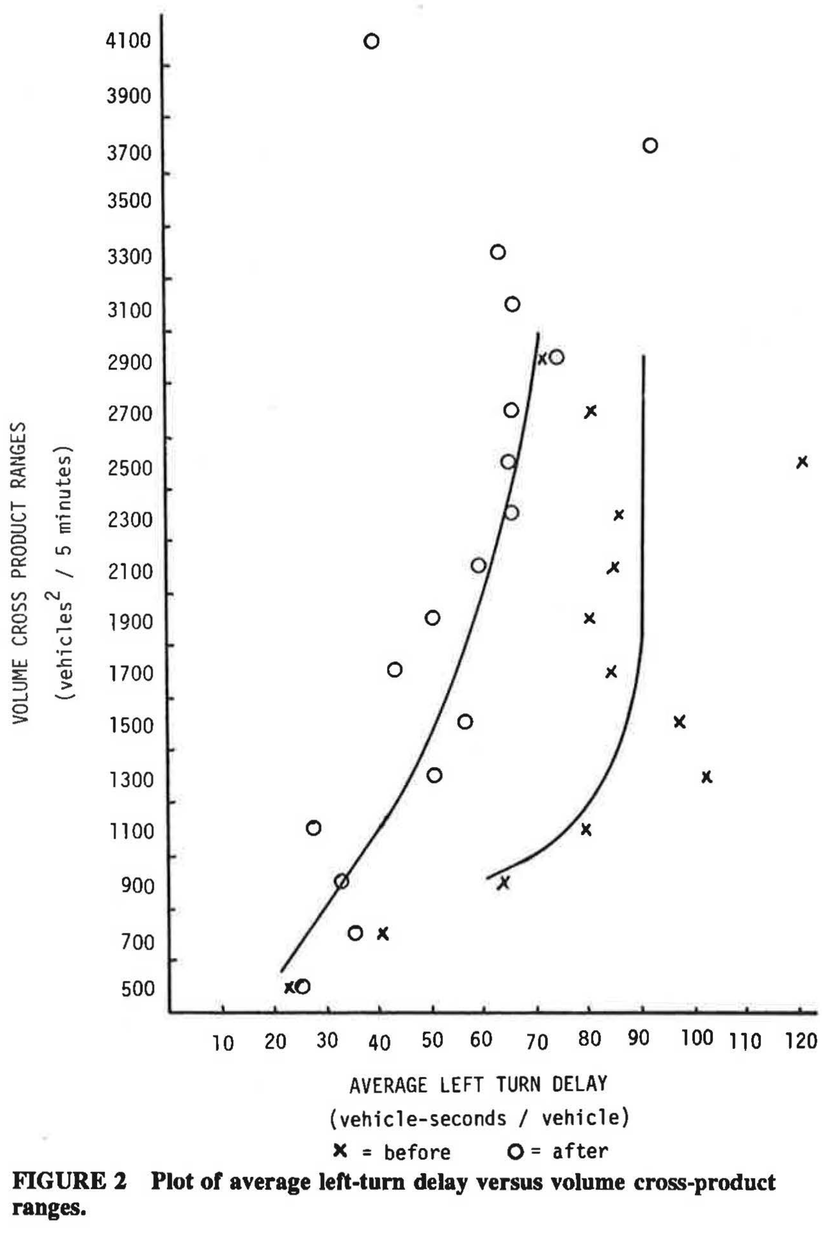


FIGURE 2 Plot of average left-turn delay versus volume cross-product ranges.

TABLE 9 PERCENT OF VEHICLES STOPPING (8-hr average)

	Before	After	Percent Difference
Southbound through	34.7	64.5	+85.7
Southbound left turn	94.9	98.0	+3.2
Northbound through	35.7	67.7	+89.8
Northbound left turn	98.4	97.6	−0.8
SB + NB through and left-turn combined	42.7	70.6	+65.4

were statistically significant. The through-traffic movements accounted for a very large proportion of the approach volume (86.3 percent of total observed volume in the after phase).

The slight changes in the fraction of left-turn vehicles that stopped were minimal. At least 95 percent of left-turn vehicles stopped in both the before and after phases.

For all movements combined, the percent of vehicles that stopped increased from 42.7 percent in the before phase to 70.6 percent in the after phase.

Economic Impact

It has been demonstrated in previous sections of this paper that vehicle delay and the number of vehicles stopping were both greatly increased. These impacts result in greatly increased costs for the roadway user and the public in terms of increased vehicle operating cost, increased travel time, and vehicle emissions. These costs can be estimated based on: (a) the increased number of vehicles that must decelerate and accelerate due to the increased percentage of vehicles that stop; and (b) the increased stopped delay (vehicle hours of idling time). An estimate of these costs is described next.

Costs Due to Additional Number of Stopping Vehicles

An estimate of additional vehicle operating costs was performed by using procedures described in *A Manual on User Benefit Analysis of Highway and Bus Transit Improvements, 1977 (6)*. Unit costs presented in this report were updated to 1985 values, and a correction was made for vehicle fleet fuel consumption improvements.

From the time-lapse photography it was known that during the 8 hr that were filmed the number of vehicles that stopped increased by 6,868 vehicles. Expanding this number to a 24-hr period yielded 9,444 vehicles per day. For simplicity, it was assumed that (a) all of these vehicles were passenger cars; (b) they all underwent a speed change cycle from 40 to 0 to 40 mph; and (c) vehicles that did not stop did not go through a speed change cycle at all. These assumptions caused the actual increase in costs to be understated.

The additional vehicle operating cost due to a speed change cycle from 40 to 0 to 40 mph (updated to a 1985 value) is $31.74 per 1,000 speed change cycles. Therefore, the additional vehicle operating cost on 44th Street was

$$\$31.74/1{,}000 \text{ vehicles} \times 9{,}444 \text{ vehicles} = \$299.77/\text{day}$$

This equals $109,416/year.

The value of time is related to the activity pursued and the length of time involved. It is difficult to quantify such a subjective cost. Previous studies have assigned values to low, medium, and high time savings, which should logically be applicable to losses of time as well *(6)*. The medium time savings value per traveler-hour for average trips was chosen as a reasonable estimate. This was $1.80/hr at 1975 prices *(6)*. Updated to 1985, the value is $3.66/hr per person.

Average vehicle occupancy was assumed to be 1.56. This is the same value used in 1977 after an FHWA survey indicated it to have held true since the late 1960s *(7)*. This increases the cost to $5.71/hr per vehicle.

Travel time losses were calculated from a base value of 4.42 hr/1,000 speed change cycles *(8)*. Multiplied by 9,444 additional vehicles stopping, the result was 41.74 additional hours of travel time per day. At $5.71/hr, this amounts to a travel time cost of $238.34/day or $86,992.98/year.

Data on increased vehicle emissions due to speed change cycles are provided by Dale *(8)*. Unit costs of the pollutants are also available *(9)*.

An additional 46 lb of carbon monoxide, 2.1 lb of hydrocarbons, and 2.4 lb of nitrogen oxide are generated for every 1,000 stopping vehicles. By applying the unit costs and multiplying by 9,444 vehicles, the additional emission costs were calculated to be $6.60/day or $2,408.60/year. The increased costs due to the additional number of stopping vehicles are summarized in Table 10.

TABLE 10 ADDITIONAL COST FOR THE ROADWAY USER AND THE PUBLIC

	Additional Cost per Day Due to Additional No. of Stopping Vehicles ($)	Additional Cost per Day Due to Increased Stopped Delay ($)	Total Additional Cost per Day ($)	Total Additional Cost per Year ($)
Vehicle operating cost	299.77	43.12	342.89	125,155
Travel time cost	238.34	501.91	740.25	270,191
Emission cost	6.60	2.28	8.88	3,241
Total	544.71	547.31	1,092.02	398,587

Costs Due to Increased Stopped Delay

A Manual on User Benefit Analysis of Highway and Bus-Transit Improvements (6) was also used to estimate these additional vehicle operating costs. Once again, unit costs were updated to 1985 values and a correction was made for improved fuel economy. The 1985 rate was $490.60 per 1,000 vehicle-hours of idling. Multiplied by the additional 87.9 vehicle-hours of stopped delay, this yielded an increased vehicle operating cost of $43.12/day.

The value of a person's time is the largest component of the costs due to stopped delay. Using the 87.9 additional vehicle-hours of delay per day and a value of $5.71/hr per vehicle yields a daily travel time cost of $501.91.

An additional 434 lb of carbon monoxide, 14.1 lb of hydrocarbons, and 4.4 lb of nitrogen oxides are generated by 87.9 vehicle-hours of idling. The emission cost of these pollutants is $2.28/day.

All additional costs due to idling and stopping are given in Table 10 in terms of cost per day and cost per year. It is emphasized that the values in Table 10 are the increased costs for vehicles on 44th Street only. If increased costs for Thomas Road were also considered, the costs would be much higher.

CONCLUSIONS

1. Volumes of both left-turn and through movements for north and south approaches increased in the after phase. The increase in left-turn volumes was found to be statistically significant; the increase in through volumes was not.
2. The increase in southbound left-turn volume, expressed as a percentage of total southbound volume, was significant at the 95 percent level of confidence.
3. The increase in northbound left-turn volume was not significant when expressed as a percentage of total northbound volume.
4. Average delay to southbound through vehicles more than quadrupled in the after phase. Average delay to northbound through vehicles more than tripled in the after phase.
5. Average delays to left-turn vehicles decreased to 82 percent of the before values. The decrease was not found to be statistically significant. Delay decreased even though left-turn volume increased.
6. There was a minimum net increase in total delay (on the north and south approaches) of 87.9 vehicle-hours per day. The net decreases in left-turn delay were only a fraction of the net increases in through delay.
7. The increased vehicle operating, travel time, and emission costs due to the net increase in delay were at least $547/day or $199,655/year. Additional costs due to the increased number of stopping vehicles were $545/day or $198,819/year. The combined costs were $1,092/day and $398,587/year.
8. The percentage of through vehicles that stopped on 44th Street increased significantly (from 35 to 66 percent).
9. Longer cycle lengths and inefficient use of green time increased the number of stopping vehicles and vehicle delay.
10. The loss of progression contributed to the problem of inefficient use of through-green time.
11. The efficiency of through movement operations was impaired by the addition of the exclusive phase. The improvements in processing left-turn vehicles were obtained at the expense of inconveniencing the through movement.

ACKNOWLEDGMENT

Technical and financial assistance were provided for this study by the city of Phoenix.

REFERENCES

1. Institute of Transportation Engineers—Colorado-Wyoming Section. *A Study of the Use of Warrants for the Installation of Left-Turn Phasing at Signalized Intersections.* Washington, D.C., March 1985.
2. K. R. Agent and R. C. Deen. *Warrants for Left-Turn Signal Phasing.* Division of Research, Bureau of Highways, Department of Transportation, Commonwealth of Kentucky, Lexington, Oct. 1978.
3. R. M. Mohle and T. Rorabaugh. *Left Turn Phasing, Vol. 4: A Study of Clearance Intervals, Flashing Operation, and Left-Turn Phasing at Traffic Signals.* Report FHWA-RD-78-49. FHWA, U.S. Department of Transportation, May 1980.
4. D. L. Warren. Accident Analysis of Left-Turn Phasing. *Public Roads,* Vol. 48, No. 4, March 1985.
5. J. E. Upchurch. Guidelines for Selecting Type of Left Turn Phasing. In *Transportation Research Record 1069,* TRB, National Research Council, Washington, D.C., 1986, pp. 30–38.
6. *A Manual on User Benefit Analysis of Highway and Bus-Transit Improvements.* AASHTO, Washington, D.C., 1978.
7. T. C. Thomas and G. I. Thompson. Value of Time Saved by Trip Purpose. In *Highway Research Record 369,* TRB, National Research Council, Washington, D.C., 1971, pp. 104–117.
8. C. W. Dale. *Procedure for Estimating Highway User Costs, Fuel Consumption and Air Pollution.* Office of Traffic Operations, FHWA, U.S. Department of Transportation, March 1980.
9. K. A. Small. Estimating the Air Pollution Costs of Transport Modes. *Journal of Transport Economics and Policy,* Vol. II, No. 2, May 1977, pp. 109–132.

DISCUSSION

E. C. P. Chang

Texas Transportation Institute, Texas A&M University System, College Station, Tex. 77843-3135.

Left-turn control strategies have important impacts on the signal capacity, traffic operations, and safety design of signalized intersections. Selecting the proper left-turn signal phasing can improve the level of service, decrease intersection delay, and reduce left-turn-related accidents. Various jurisdictions frequently have to determine which left-turn treatments are more effective for daily traffic operations. Three left-turn signal phasings are commonly used with the green arrow or circular green indications. These control strategies include the "permissive or permitted," "exclusive or protected," and "exclusive/permissive or protected/permitted" phases for different signalized left-turn treatments. However, there are currently no standard guidelines in the United States for determining which left-turn phasing treatment is best at a particular intersection.

Permissive versus exclusive phasings have been discussed in many past studies. However, few studies have examined permissive versus exclusive/permissive treatment. Many practicing traffic engineers have been reluctant to convert from permissive to exclusive/permissive phasings for three reasons. First, they are likely to reduce arterial through-green times for progression. Second, the possible delay increase may reduce total intersection capacity. Third, no good signal timing methods are currently available to calculate exclusive/permissive green splits and provide capacity evaluation. Despite these potential disbenefits, the uses of exclusive phases and alternative phase sequences have been proved to be successful in many signalized locations. Overall, implementation of the exclusive left-turn treatment can significantly reduce the possibility of severe left-turn accidents when a large percentage of left-turn traffic exists at a signalized intersection.

Described in this paper is a field experiment study that investigated the conversion from permissive to exclusive/permissive operation. The study compared two signal treatments and evaluated the before-and-after performance. A signalized intersection in Phoenix, Arizona, was converted from pretimed to semi-actuated operation. Traffic volume and delay data were obtained and processed manually through time-lapse photography. Stopped delay and traffic characteristics were identified for the arterial left-turn and through movements. Finally, performance evaluations were summarized for the specific study intervals. In this study

1. A statistically significant amount of volume increase in left-turn movements was observed in the "after" study.
2. A significant increase in average arterial through delay was noted when the signal control changed from the "before" permissive left turn to the "after" exclusive/permissive left-turn treatment.
3. A decrease in left-turn delay in the after study was observed even though the overall volume had increased. However, the net increases in arterial through delay were far greater than the net decreases in left-turn delay after the exclusive/permissive phase was implemented.

The intent of this study was to evaluate the differences in operational performance between permissive and exclusive/permissive operations. However, there are three major concerns about the results of this field evaluation. First, because the study team was not able to control the development of traffic signal timing plans for after comparisons, the results may not be suitable for drawing general conclusions among the permissive, exclusive, or exclusive/permissive arterial operations. Second, because the researchers were not able to design desirable signal timing plans to account for the possible effects, some of the observed findings may actually originate from the fact that these signal timing parameters may not be set properly for exclusive/permissive operations. Third, the statements concerning the loss of progression as a result of the use of the exclusive phase are somewhat misleading.

Because of the preceding concerns, three additional comments are recommended:

1. To use the exclusive or exclusive/permissive left turn effectively, the signal timing plans, and especially the arterial phase sequences, have to be provided properly in order to allow maximum arterial progression and yet maintain minimum stops and total delay.
2. The timing design of exclusive/permissive or permissive/exclusive phasings relies primarily on how to provide short but sufficient green time for the required protected left-turn phase movements.
3. Effective signal system operation requires efficient coordination between arterial capacity analysis and signal timing optimization.

When designed and implemented properly, exclusive or exclusive/permissive phases can effectively clear the arterial left-turn traffic in advance of the arriving progression traffic, thereby increasing effective signal capacity and improving operational safety. Therefore, two operational considerations are needed for a fair before-and-after comparison. First, the revised arterial timing design is needed to generate effective, coordinated progression offsets for exclusive/permissive signal operations. Second, accurate signal capacity analyses are also required to calculate efficient amounts of green splits for the exclusive portion of the total left-turn phase. Essentially, three basic design questions have to be answered:

1. What amount of effective green time can be allocated for the protected portion of the left-turn phase without having to take the opposing through green needed for coordinated arterial progression?
2. How should the permitted left-turn saturation flow be accounted for in the permitted left-turn phase to reflect the equivalent added signal capacity in the arterial directions because of the increased arterial through-green time in the permissive phase?
3. How much should the arterial traffic be adjusted to consider the increased arrival traffic in the arterial directions due to the "platooned" traffic from the arterial signal progression effects?

This study confirmed that signal timing design for exclusive/permissive or permissive/exclusive left-turn operations is an important yet complicated process. Field performance measurements are extremely susceptible to the way signal timing plans are implemented and perceived by motorists. Normally, arterial travel time and stopped delay can be reduced by carefully timing traffic signal systems for efficient progression operations. Each of the permissive, exclusive, or exclusive/permissive left-turn signal treatments may introduce operational problems to the arterial system if they have not been timed properly for coordinated system operation. Therefore, the impacts of timing plans on signalized intersection delay should be thoroughly examined before any field implementation can be proved to be successful. Simulation studies or field experiments should not only be performed at individual intersections, but the resultant arterial progression should also be carefully investigated. In this way, more comparative before-and-after study results can be used to examine different traffic signal control strategies before implementing signal timing plans.

AUTHORS' CLOSURE

Chang's comments are greatly appreciated; they stimulate much needed discussion on this important topic.

As indicated in the paper, the study team did not have control over the signal timing plans used in the "after" portion of the study period. Timing plans such as cycle length and G/C ratio for each movement were factors that the study team simply had to accept. Whether or not the signal timing parameters in the after phase were good or poor is simply conjecture at this point.

We generally agree with the following comments made by Chang.

1. To use the exclusive or exclusive/permissive left turn effectively, the signal timing plans, and especially the arterial phase sequences, have to be provided properly in order to allow maximum arterial progression and yet maintain minimum stops and total delay.
2. The timing design of exclusive/permissive or permissive/

exclusive phasings relies primarily on how to provide short but sufficient green time for the required protected left-turn phase movements.

3. Effective signal system operation requires efficient coordination between arterial capacity analysis and signal timing optimization.

Chang states: "When designed and implemented properly, exclusive or exclusive/permissive phases can effectively clear the arterial left-turn traffic in advance of the arriving progression traffic, thereby increasing effective signal capacity and improving operational safety." This is fairly easy to accomplish on a single arterial street. However, it is more difficult to accomplish in a connected network of arterial streets. As Chang points out, "Each of the permissive, exclusive, or exclusive/permissive left-turn signal treatments may introduce operational problems to the arterial system if they have not been timed properly for coordinated system operation." We agree.

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

Operational Analysis of Exclusive Left-Turn Lanes with Protected/Permitted Phasing

JAMES A. BONNESON AND PATRICK T. MCCOY

With the release of the *1985 Highway Capacity Manual*, a new procedure for analyzing signalized intersections has been introduced. One of the major differences between the 1965 and the 1985 manuals is in the area of left-turn capacity. A general methodology for the analysis of signalized intersections, particularly left-turn operations, is described in the 1985 Manual. Unfortunately, with regard to exclusive left-turn lanes with protected/permitted phasing, the sample calculations provided do not appear to explicitly follow the general methodology. Moreover, the sample calculations introduce many new concepts that are not included in the discussion of the methodology. Calculation 3 in Chapter 9 of the manual is reexamined in this paper. In particular, the left-turn lane groups with protected/permitted phasing are reanalyzed according to the general methodology, but issue is taken with some of the "new" concepts introduced within Calculation 3. On the basis of the findings reported in this paper, it appears that there is a need for some revision of Chapter 9 of the *1985 Highway Capacity Manual*, particularly with regard to the analysis of left-turn lane groups with protected/permitted phasing.

Chapter 9 of the *1985 Highway Capacity Manual* (HCM) *(1)* contains procedures for evaluating the capacity and level of service of signalized intersections. The operational-analysis methodology presented in the HCM accounts for the effect of left-turn movements based on the manner in which they are accommodated. In the case of left turns made from an exclusive left-turn lane controlled by protected/permitted phasing, the HCM recommends an iterative procedure, which is shown in Figure 1. In this procedure, all left turns are initially assumed to occur in the protected phase. If this assumption results in volume-to-capacity ratios that are too high, a portion of the left turns, up to the capacity of the permitted phase, is assigned to the permitted phase, and the saturation-flow-rate and capacity-analysis modules are repeated. The portion of left-turns assigned to the permitted phase is increased on successive iterations until either acceptable volume-to-capacity ratios are obtained or the capacity of the permitted phase is reached.

Unfortunately, only a general description of this procedure is given in the HCM. Also, the sample calculations presented in the HCM do not correctly illustrate the procedure as it is described. Consequently, the generality of its description and the inconsistency between this description and the sample calculation illustrating its use have been sources of confusion to HCM users.

J. A. Bonneson, Traffic Operations Program, Texas Transportation Institute, Texas A&M University System, College Station, Tex. 77843. P. T. McCoy, Civil Engineering Department, University of Nebraska, Lincoln, Lincoln, Neb. 68588.

In an effort to eliminate this confusion, the operational-analysis procedure presented in the HCM for evaluating the capacity and level of service of exclusive left-turn lanes controlled by protected/permitted phasing is reviewed in this paper, and revisions to the procedure are suggested to make it consistent with other procedures in the HCM. The revised procedure is presented within the context of a reanalysis of Calculation 3, which is the sample calculation used in the HCM to illustrate the operational analysis of exclusive left-turn lanes with protected/permitted phasing. The reanalysis of Calculation 3 is presented in the first section of this paper, which includes explanations of the revisions made to the procedure presented in the HCM. The second section includes a summary of the revised procedure recommended. The solution of Calculation 3 is compared with the solution of Calculation 3 presented in the HCM.

REANALYSIS OF CALCULATION 3

Calculation 3, which begins on page 9-50 of the HCM, is the operational analysis of a multiphase-actuated signal located at the intersection of Fifth Avenue and 12th Street. The input worksheet showing the geometric, traffic, and signalization conditions at this intersection is shown in Figure 2. Exclusive left-turn lanes are provided on all four approaches to the intersection. Protected/permitted phasing is provided for the left-turns from the north-south street (Fifth Avenue), and permitted phasing is provided for the left turns from the east-west street (12th Street).

In this reanalysis, the procedure suggested in the HCM is followed except where revisions are noted. All pertinent worksheets are completed and shown in this reanalysis. However, the discussion focuses only on those points in the solution where revisions to the procedure are made. Although this paper is concerned with just the operational analysis of exclusive left-turn lanes with protected/permitted phasing, all worksheets are completed for the entire intersection to better illustrate the consequences of the revisions. A discussion of the reanalysis of Calculation 3 with respect to each of the modules in the operational-analysis procedure follows.

Input and Volume Adjustment Modules

The input and volume adjustment modules are performed in the same way as they were in the original analysis. The worksheets for these modules are shown in Figures 2 and 3. They are

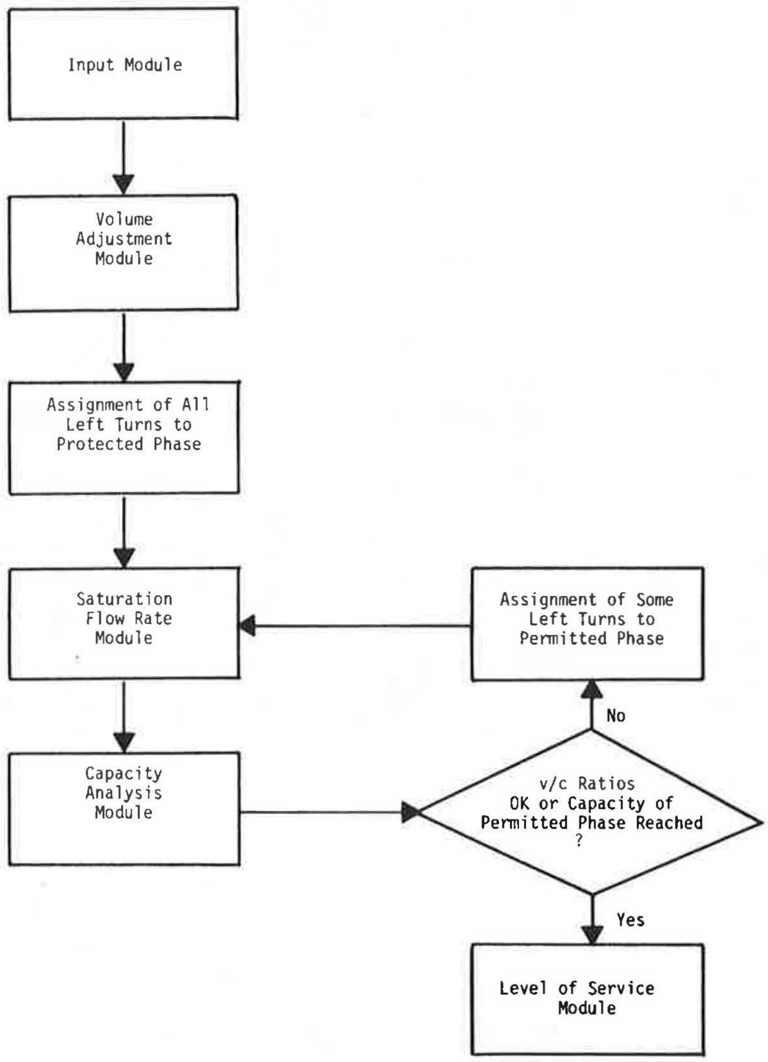


FIGURE 1 Iterative procedure for operational analysis of exclusive left-turn lanes with protected/permitted phasing.

identical to those shown in the HCM and are presented here only for convenience.

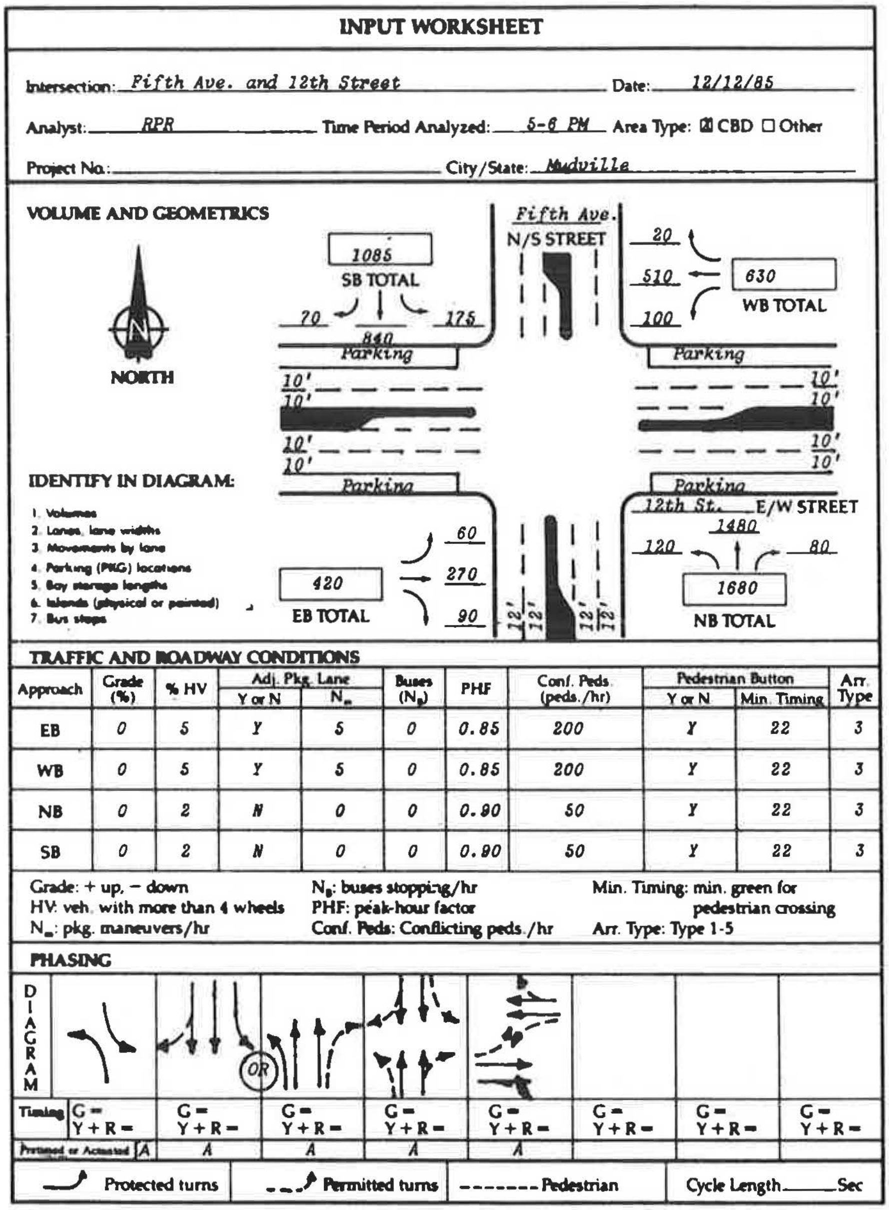
Saturation Flow Rate Module

The saturation flow adjustment worksheet is shown in Figure 4. The adjustment factors used are identical to those used in the HCM with one exception: the left-turn adjustment factors for the eastbound (EB) and westbound (WB) permitted left-turns have been modified (i.e., EB: 0.29, not 0.31 and WB: 0.46, not 0.48). The reason for this deviation is explained by the iterative nature of the method for computing left-turn adjustment factors for permitted left turns. In other words, for those situations where the signal timing is not known or where the signal is actuated, as in this case, the corresponding phase durations must be initially estimated and then solved for iteratively. Ultimately, the assumed signal timings will converge to reasonable values, and the result will most accurately reflect the intersection's operation.

In the original analysis of Calculation 3, a 90-sec cycle and a 18.5-sec phase duration were initially assumed for the calculation of the eastbound and westbound left-turn adjustment factors. This represents a good starting solution. But if the analyst had iterated through the analysis procedure, better estimates of these times would have been obtained as shown in Figure 5. Thus, Calculation 3 as presented in the HCM illustrates only the first iteration of the analysis process, whereas the results shown in Figures 4 and 5 are representative of the last iteration and hence the saturation flow rates shown should be more accurate than those shown in Figure 9-28 of the HCM.

It should also be noted that some of the saturation flow rates for other movements differ by 1 or 2 percent. This amount is negligible and can be attributed to the effects of rounding during the analysis process.

For the purposes of comparison between this analysis and that presented in the HCM, a cycle length of 119 sec is used for all subsequent analysis steps. This approach highlights deviations resulting from the analysis process rather than those attributable to different cycle lengths.



INPUT WORKSHEET

Intersection: Fifth Ave. and 12th Street Date: 12/12/85

Analyst: RPR Time Period Analyzed: 5-6 PM Area Type: ☒ CBD ☐ Other

Project No.: City/State: Mudville

TRAFFIC AND ROADWAY CONDITIONS

Approach	Grade (%)	% HV	Adj. Pkg. Lane Y or N	Adj. Pkg. Lane N_m	Buses (N_B)	PHF	Conf. Peds. (peds./hr)	Pedestrian Button Y or N	Pedestrian Button Min. Timing	Arr. Type
EB	0	5	Y	5	0	0.85	200	Y	22	3
WB	0	5	Y	5	0	0.85	200	Y	22	3
NB	0	2	N	0	0	0.90	50	Y	22	3
SB	0	2	N	0	0	0.90	50	Y	22	3

Grade: + up, − down
HV: veh. with more than 4 wheels
N_m: pkg. maneuvers/hr
N_B: buses stopping/hr
PHF: peak-hour factor
Conf. Peds: Conflicting peds./hr
Min. Timing: min. green for pedestrian crossing
Arr. Type: Type 1-5

PHASING

Timing	G = Y + R =	G = Y + R =	G = Y + R =	G = Y + R =	G = Y + R =	G = Y + R =	G = Y + R =	G = Y + R =
Pretimed or Actuated	A	A	A	A	A			

Protected turns — Permitted turns — ------- Pedestrian — Cycle Length ____ Sec

FIGURE 2 Input worksheet for Calculation 3.

Capacity Analysis Module

The capacity analysis worksheet is shown in Figure 6. Given the phasing plan shown on the input worksheet in Figure 2, the combinations of critical lane groups are found according to the following rule:

$$\begin{bmatrix} \text{EB LT or TH/RT} \\ \text{or} \\ \text{WB LT or TH/RT} \end{bmatrix} + \begin{bmatrix} \text{NB LT + SB TH/RT} \\ \text{or} \\ \text{SB LT + NB TH/RT} \end{bmatrix}$$

Thus, the sum of critical flow ratios results from the combination: WB TH/RT + SB LT + NB TH/RT = 0.24 + 0.09 + 0.57 = 0.90. This represents the percentage of green time needed to adequately serve intersection traffic during the analysis hour.

Average Cycle Length and Lost Time

Using Equation II.9-1 in the HCM, the cycle length is computed as follows:

$$C = LX_c/[X_c - \sum_i (v/s)_{ci}] \qquad (1)$$

where

C = cycle length, in seconds;
L = total lost time per cycle, in seconds;
X_c = critical v/c ratio for the intersection; and
$(v/s)_{ci}$ = sum of critical flow ratios.

VOLUME ADJUSTMENT WORKSHEET										
① Appr.	② Mvt.	③ Mvt. Volume (vph)	④ Peak Hour Factor PHF	⑤ Flow Rate v_p (vph) ③÷④	⑥ Lane Group	⑦ Flow rate in Lane Group v_g (vph)	⑧ Number of Lanes N	⑨ Lane Utilization Factor U Table 9-4	⑩ Adj. Flow v (vph) ⑦×⑨	⑪ Prop. of LT or RT P_{LT} or P_{RT}
EB	LT	60	0.85	71		71	1	1.00	71	1.0 LT
	TH	270	0.85	318		424	2	1.05	445	0.25 RT
	RT	90	0.85	106						
WB	LT	100	0.85	118		118	1	1.00	118	1.0 LT
	TH	510	0.85	600		624	2	1.05	655	0.04 RT
	RT	20	0.85	24						
NB	LT	120	0.90	133		133	1	1.00	133	1.0 LT
	TH	1480	0.90	1644		1733	2	1.05	1820	0.05 RT
	RT	80	0.90	89						
SB	LT	175	0.90	194		194	1	1.00	194	1.0 LT
	TH	840	0.90	933		1011	2	1.05	1062	0.08 RT
	RT	70	0.90	78						

FIGURE 3 Volume adjustment worksheet for Calculation 3.

According to the HCM, the X_c value for a fully actuated signal can be estimated at 0.95. This estimate is based on the additional assumption that actuated intersections operate efficiently with respect to the allocation of green time. Intuitively, this approach is reasonable and should provide a good approximation of the average signal timing during the analysis hour.

At this point, some discussion is necessary for the determination of total intersection lost time. The HCM states that total lost time per cycle for this intersection and phasing combination is 6.0 sec. This value represents two 3.0-sec increments of lost time corresponding to the two through phases and assumes there is continuous utilization of the time element occurring between overlapped phases. Although this argument appears at first to have validity, it is incorrect. By definition, lost time is the time lost due to start-up, delay, and intersection clearance (totaling approximately 3.0 sec) that is experienced by each critical lane group. Hence, it is experienced by all three critical lane groups associated with Calculation 3 for a total of 9.0 sec of lost time.

Based on the assumption of X_c equal to 0.95 and a total lost time of 6.0 sec, the average cycle length was estimated to be 118.8 sec in the original Calculation 3 analysis. Using this cycle length, the effective green times were estimated by proportionally allocating the total cycle length to the critical lane groups as given in Table 1. However, because of the initially incorrect assumption of total lost time, the sum of the phase lengths given in Table 1 is greater than the cycle length by 3.0 sec, the amount by which the lost time was underestimated.

Assuming that total intersection lost time is 9.0 sec, a more realistic average cycle length can be estimated using Equation 1. For a 9.0-sec lost time and the given X_c of 0.95, the average cycle length is calculated to be 171 sec. This represents a 44

SATURATION FLOW ADJUSTMENT WORKSHEET												
LANE GROUPS				ADJUSTMENT FACTORS								
① Appr.	② Lane Group Movements	③ Ideal Sat. Flow (pcphgpl)	④ No. of Lanes N	⑤ Lane Width f_w Table 9-5	⑥ Heavy Veh f_{HV} Table 9-6	⑦ Grade f_g Table 9-7	⑧ Pkg. f_p Table 9-8	⑨ Bus Blockage f_{bb} Table 9-9	⑩ Area Type f_a Table 9-10	⑪ Right Turn f_{RT} Table 9-11	⑫ Left Turn f_{LT} Table 9-12	⑬ Adj. Sat. Flow Rate s (vphg)
		1800	1	.930	.975	1.0	1.0	1.0	.90	1.0	.291	429
EB		1800	1	.930	.975	1.0	.935	1.0	.90	.942	1.0	2590
		1800	1	.930	.975	1.0	1.0	1.0	.90	1.0	.457	675
WB		1800	2	.930	.975	1.0	.935	1.0	.90	.990	1.0	2734
		1800	1	.99	1.0	1.0	1.0	1.0	.90	1.0	.95	1524
NB		1800	2	.99	1.0	1.0	1.0	1.0	.90	.99	1.0	3180
		1800	1	.99	1.0	1.0	1.0	1.0	.90	1.0	.95	1524
SB		1800	2	.99	1.0	1.0	1.0	1.0	.90	.99	1.0	3166

FIGURE 4 Saturation flow adjustment worksheet for Calculation 3.

percent increase in cycle length over the original estimate of 118.8 sec. Moreover, this illustrates the sensitivity of Equation 1 to estimates of L, $\sum_i (v/s)_{ci}$, and X_c.

But, as previously mentioned, a 119-sec cycle length is used for all subsequent steps in this reanalysis in order to provide a more direct comparison between it and the original analysis of Calculation 3. Hence, to maintain the equality in Equation 1, X_c must be calculated. Given a 119-sec cycle length, a $\sum_i (v/s)_{ci}$ of 0.90, and a 9.0-sec lost time, X_c is found to be 0.97.

The signal timing plan for a 119-sec cycle and a 9.0-sec lost time is given in Table 2. In this case, as expected, the sum of the phase lengths is equal to the cycle length.

Left-Turn Capacity

During the iteration process, the amount of left-turn volume assigned to the protected portion of the protected/permitted left-turn phase is reduced, if possible. This reduction is a function of the theoretical capacity of the permitted phase portion. As specified in the HCM (see Step 10, p. 9-30), the

SUPPLEMENTAL WORKSHEET FOR LEFT-TURN ADJUSTMENT FACTOR, f_{LT}				
INPUT VARIABLES	EB	WB	NB	SB
Cycle Length, C (sec)	119.0	119.0	119.0	119.0
Effective Green, g (sec)	29.3	29.3	80.7	70.0
Number of Lanes, N	1.0	1.0	1.0	1.0
Total Approach Flow Rate, v_a (vph)	494	741	1866	1205
Mainline Flow Rate, v_M (vph)	423	623	1733	1011
Left-Turn Flow Rate, v_{LT} (vph)	70.6	117.6	133	194
Proportion of LT, P_{LT}	1.0	1.0	1.0	1.0
Opposing Lanes, N_o	2.0	2.0	2.0	2.0
Opposing Flow Rate, v_o (vph)	623	423	1011 / 1062	1733 / 1820 (a)
Prop. of LT in Opp. Vol., P_{LTO}	0.0	0.0	0.0	0.0
COMPUTATIONS	EB	WB	NB	SB
$S_{op} = \dfrac{1800 N_o}{1 + P_{LTO}\left[\dfrac{400 + v_M}{1400 - v_M}\right]}$	3600	3600	3600 / 3166	3600 / 3180 (b)
$Y_o = v_o / S_{op}$	0.173	0.118	0.281 / 0.335	0.481 / 0.572
$g_u = (g - CY_o) / (1 - Y_o)$	10.5	17.3	65.7 / 61.4	24.6 / 4.4
$f_s = (875 - 0.625 v_o) / 1000$	0.485	0.610		
$P_L = P_{LT}\left[1 + \dfrac{(N-1)g}{f_s g_u + 4.5}\right]$	1.0	1.0		
$g_q = g - g_u$	18.8	12.0		
$P_T = 1 - P_L$	0.0	0.0		
$g_f = 2\dfrac{P_T}{P_L}\left[1 - P_T^{\,0.5 g_q}\right]$	0.0	0.0		
$E_L = 1800 / (1400 - v_o)$	2.32	1.84		
$f_m = \dfrac{g_f}{g} + \dfrac{g_u}{g}\left[\dfrac{1}{1 + P_L (E_L - 1)}\right] + \dfrac{2}{g}(1 + P_L)$	0.291	0.457		
$f_{LT} = (f_m + N - 1) / N$	0.291	0.457		

FIGURE 5 Supplemental worksheet for Calculation 3.

capacity of the permitted left-turn phase is calculated as the maximum of

$$C_{LT} = (1{,}400 - V_o)\,(g/C)_{PLT} \tag{2}$$

or

$$C_{LT} = 2 * 3{,}600/C \tag{3}$$

where

C_{LT} = capacity of the left-turn permitted phase, in vph;
V_o = opposing through plus right-turn flow rate, in vph; and
$(g/C)_{PLT}$ = effective green ratio for the permitted left-turn phase.

Unfortunately, the methodology does not describe in sufficient detail the derivation of the $(g/C)_{PLT}$ ratio. Even more unfortunate is the omission of Equation 2 from the discussion of Calculation 3. The effects of this omission will be more evident in the next few paragraphs.

Further investigation of the effective-green-time term (g) in the $(g/C)_{PLT}$ ratio reveals that it is identical to the unsaturated green time that is used in the calculation of the left-turn satura-

CAPACITY ANALYSIS WORKSHEET								
LANE GROUP		③ Adj. Flow Rate v (vph)	④ Adj. Sat. Flow Rate s (vphg)	⑤ Flow Ratio v/s ③ ÷ ④	⑥ Green Ratio g/C	⑦ Lane Group Capacity c (vph) ④ × ⑥	⑧ v/c Ratio X ③ ÷ ⑦	⑨ Critical ? Lane Group
① Appr.	② Lane Group Movements							
EB		71	429	0.164	0.246 29.3 sec	106	0.669	
		445	2590	0.172	0.246 29.3 sec	637	0.698	
WB		118	675	0.174	0.246 29.3 sec	166	0.709	
		655	2734	0.240	0.246 29.3 sec	672	0.974	X
NB		0	1524	0.0	0.0	0	-	
		133	389	0.343	0.516 61.4 sec	201	0.662	
		1820	3180	0.572	0.588 70.0 sec	1869	0.974	X
SB		134	1524	0.088	0.091 10.7 sec	138	0.974	X
		60	1600	0.038	0.038 4.5 sec	60	1.00	
		1062	3166	0.335	0.678 80.7 sec	2148	0.493	

Cycle Length, C 119 sec

$\Sigma (v/s)_{ci} =$ 0.900

Lost Time Per Cycle, L 9.0 sec

$$X_c = \frac{\Sigma (v/s)_{ci} \times C}{C - L} = 0.974$$

FIGURE 6 Capacity analysis worksheet for Calculation 3.

TABLE 1 SIGNAL TIMING USING HCM CALCULATION OF LOST TIME

Movement	Critical Flow Ratio	Effective Green[a] (sec)	Lost Time (sec)	Phase Length[b] (sec)
EB/WB through	0.241	30.1	3.0	33.1
SB left-turn	0.088	11.0	3.0	14.0
NB/SB through	0.573	71.7	3.0	74.7
Total	0.902	112.8	9.0	121.8[c]

NOTE: Cycle length (C) = 118.8 sec, and critical v/c (X_c) = 0.95.
[a]Effective green (g) = (critical flow ratio)(C/X_c).
[b]Phase length (G) = g + lost time.
[c]Greater than cycle length (C = 118.8 sec).

TABLE 2 SIGNAL TIMING USING REVISED CALCULATION OF LOST TIME

Movement	Critical Flow Ratio	Effective Green[a] (sec)	Lost Time (sec)	Phase Length[b] (sec)
EB/WB through	0.24	29.3	3.0	32.3
SB left-turn	0.09	10.7	3.0	13.8
NB/SB through	0.57	70.0	3.0	72.9
Total	0.90	110.0	9.0	119.0[c]

NOTE: Cycle length (C) = 119.0 sec, and critical v/c (X_c) = 0.97.
[a]Effective green (g) = (critical flow ratio)(C/X_c).
[b]Phase length (G) = g + lost time.
[c]Equal to cycle length (C = 119.0 sec).

tion flow adjustment factor for permitted left turns. This unsaturated green time, which is calculated on the supplemental worksheet, is computed as follows:

$$g_\mu = (g - CY_o)/(1 - Y_o) \tag{4}$$

where

g_μ = portion of green not blocked by the clearing of an opposing queue of vehicles, in seconds;
g = effective green time, in seconds;
C = cycle length, in seconds; and
Y_o = flow ratio for opposing approach.

At this point some discussion is warranted on the appropriate values to use in calculating the opposing flow ratio ($Y_o = V_o/S_{op}$) on the supplemental worksheet. According to the methodology, V_o is defined as the "mainline flow rate" on the opposing approach. In Calculation 3, this value was found in Column 5 of the volume adjustment worksheet. However, because this procedure is an attempt to account for the discharge time of the longest opposing queue, it is suggested that the correct value to use in this instance would be the "adjusted flow rate" found in Column 10. This value is identical to that found in Column 5 with the exception that a lane utilization factor has been applied.

Intuitively, this approach is more reasonable for estimating queue discharge time because it would account for any imbalance in lane use. Obviously, the permitted left-turn movement cannot begin until the longest opposing queue has dissipated. If the opposing approach is observed to have unequal utilization among through or right lanes, or both, then this should be accounted for via the lane utilization factor. A review of the literature on left-turn capacity supports this argument *(2, 3)*.

In addition to using the adjusted flow rate, it is also suggested that the derivation of the saturation flow on the opposing approach (S_{op}) be reconsidered. Inspection of the equation used on the supplemental worksheet to compute S_{op} indicates that it does not consider many of the adjustment factors used in the saturation flow adjustment worksheet. For this particular example the corresponding values of S_{op} taken from the saturation flow adjustment worksheet are less than 90 percent of those calculated using the supplemental worksheet. Therefore, it appears redundant to calculate the saturation flow rate again when a more appropriate value has already been computed on the saturation flow adjustment worksheet.

The implications of using the suggested values for V_o and S_{op} instead of those recommended by the HCM are shown in Figure 5. As can be seen in the northbound and southbound columns, the variation between analysis approaches can be significant. In particular, the equation for g increases in sensitivity as the flow ratio (Y_o) nears 1.0. As a result, estimates of g_μ for the southbound left-turn differ by more than a factor of 5.

For the remainder of this discussion the values of g_μ calculated by using the suggested procedure will be employed in subsequent computations. Hence, the computations of the capacity of the left-turn permitted phase (C_{LT}) are as follows:

Northbound:

$C_{LT} = (1{,}400 - 1{,}011) * 61.4/119$
$= 201$ vph <-- Maximum value

or

$C_{LT} = 2 * 3{,}600/119$
$= 60$ vph

Southbound:

C_{LT} $(1{,}400 - 1{,}733) * 4.4/119$
$= 0$ vph

or

$C_{LT} = 2 * 3{,}600/119$
$= 60$ vph <-- Maximum value

According to the HCM, "up to" the maximum value for the permissive flow rate (C_{LT}) may be assigned to the permitted portion of the protected/permitted phase. Because the exact number of vehicles arriving during each phase portion is unique to each intersection and is a function of arrival patterns and upstream progression, the number of vehicles arriving during each phase portion can vary considerably. In the case of uniformly arriving traffic, the number of left-turn vehicles arriving during the protected and permitted phase portions would be proportional to their g/C ratios.

For this reanalysis of Calculation 3, the maximum permitted flow rate is assigned to the permitted phase volume; thus minimizing the time needed for the protected left-turn phase. This approach is assumed to be more reasonable from a minimum total delay standpoint because left-turn phases typically move fewer total vehicles than through phases. Hence, protected left-turn phase lengths are typically kept as short as possible to minimize total intersection delay.

This argument is particularly applicable to pretimed signals where the protected left-turn phase interval would be set as low as practical. On the other hand, vehicular demand at actuated intersections could extend the left-turn phase beyond the minimum required and thus the permitted left-turn phase component would not realize its total potential permitted flow rate. Also it should be noted that this approach assumes that coordination for the left-turn movement is not provided because this is the most common situation.

Once the capacity of the permitted portion of the protected/permitted phase has been calculated, the left-turn volume associated with this lane group can be distributed among the appropriate phase intervals. The capacity of each phase interval is calculated as follows:

Northbound:

NB $\text{Left}_{PERM} = 201$ vph
NB $\text{Left}_{PROT} = 133 - 201 = 0$ vph

Southbound:

SB $\text{Left}_{PERM} = 60$ vph
SB $\text{Left}_{PROT} = 194 - 60 = 134$ vph

One interesting outcome from the preceding calculations is that it now appears that the permitted portion of the northbound left-turn phase has sufficient capacity to adequately serve the left-turn volume. In other words, it appears that the protected portion of the northbound left-turn phase is not necessary. As alluded to at the beginning of the previous section, it now becomes apparent that the omission of Equation 2 in calculating permitted left-turn capacity can have a significant impact on the analysis process. For Calculation 3, it shows that protection for the northbound left-turn is not warranted.

Once the protected and permitted left-turn phase volumes have been calculated, the capacity analysis worksheet can be completed by using the appropriate HCM methodology. The completed worksheet is shown in Figure 6.

Level of Service Module

In this module pertinent values from the capacity analysis worksheet shown in Figure 6 are carried forward and entered on the level of service worksheet shown in Figure 7. From these values, estimates of group delay are calculated and averaged for each approach and the intersection as a whole. Ultimately, these delays are translated into levels of service that describe the quality of traffic flow associated with each group, approach, and intersection.

With respect to calculating delay for protected/permitted left-turn phases with exclusive lanes, the HCM (*1*, p. 9-56) suggests that total delay for the left-turn lane group can be estimated by using approximations for the *g*/*C* and *v*/*c* ratios. However, these assumptions are gross estimates and can result in delays that are totally unreasonable.

The difficulty encountered when calculating the delay for left-turn movements with protected/permitted phasing arises from the variation in saturation flows during one signal cycle. This situation is shown in Figure 8 for the southbound left-turn movement. As shown in this figure, the southbound left-turn has two unique saturation flows: one during its designated left-turn phase and the other representing sneaker activity at the end of the through phase. By comparison, the protected left-turn phase for the northbound left-turn was eliminated because of ample time during the through phase for filtering left-turn operations. Hence, this movement has only one saturation flow rate.

The uniform delay incurred by left-turn vehicles can be found by calculating the area under the queue-departure diagram (shaded area) shown in Figure 8. Individual delay components can be separately calculated as that area immediately preceding the particular phase portion (i.e., the protected and permitted phase portions). For this particular example, the

LEVEL-OF-SERVICE WORKSHEET

Lane Group		First Term Delay				Second Term Delay				Total Delay & LOS		
① Appr.	② Lane Group Movements	③ v/c Ratio X	④ Green Ratio g/C	⑤ Cycle Length C (sec)	⑥ Delay d_1 (sec/veh)	⑦ Lane Group Capacity c (vph)	⑧ Delay d_2 (sec/veh)	⑨ Progression Factor PF Table 9-13	⑩ Lane Group Delay (sec/veh) (⑥+⑧) × ⑨	⑪ Lane Group LOS Table 9-1	⑫ Approach Delay (sec/veh)	⑬ Appr. LOS Table 9-1
EB		0.669	0.246	119	30.8	106	9.9	0.85	34.6	D	29.2	D
		0.698	0.246	119	31.0	637	2.3	0.85	28.4	D		
WB		0.709	0.246	119	31.1	166	8.8	0.85	34.0	D	44.7	E
		0.974	0.246	119	33.8	672	21.1	0.85	46.6	E		
NB		0.662		119	16.3	201	5.5	1.00	21.8	C	24.5	C
		0.974	0.588	119	17.9	1869	11.2	0.85	24.8	C		
SB		0.978		119	32.4	199	43.0	1.00	75.4	F	16.8	C
		0.493	0.678	119	7.0	2148	0.2	0.85	6.1	B		

Intersection Delay 26.4 sec/veh Intersection LOS D (Table 9-1)

FIGURE 7 Level-of-service worksheet for Calculation 3.

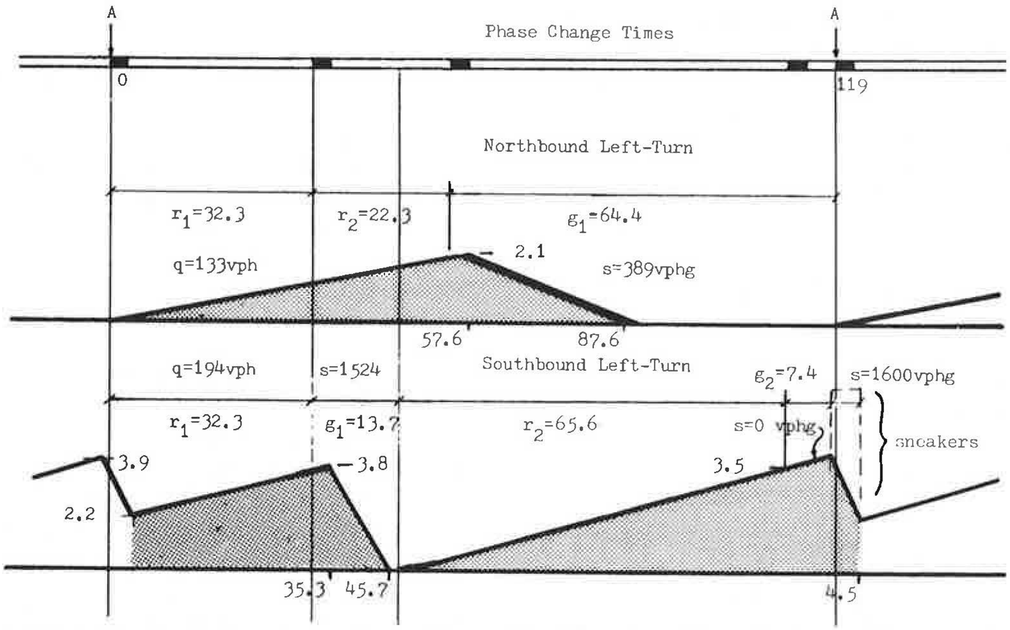


FIGURE 8 Queue departure patterns for northbound and southbound left-turn movements.

southbound left-turn uniform delay was found to incur 32.4 sec per vehicle (14.9 protected, 17.5 permitted) (see Figure 7).

The second term of delay is intended to account for the effects of random arrivals with regard to their creating overflow from one signal cycle to the next. Because this delay is based on overall cycle efficiency, it must account for left-turn operations during all phases that service left-turning vehicles. This is accomplished by calculating the combined protected and permitted capacity for the left-turn movement. As shown in Figure 7, the combined capacity for the southbound left-turn phase components is 199 vehicles per hour and results in an X ratio of 0.978 (= 194/199). Using this X ratio, the overflow delay can be calculated as 43 sec per vehicle.

As shown in Figure 9-31 of the HCM, a progression factor of 1.0 was used for the eastbound and westbound left-turn lane groups in Calculation 3. Although the use of this factor is a subjective determination by the analyst (based on first-hand knowledge of vehicular arrivals), it appears that if Table 9-13 of the HCM was followed explicitly, a factor of 0.85 would be recommended here. In particular, one of the notes accompanying this table states: "All LT^c. This category refers to exclusive LT lane groups with protected phasing only. When LT's are included . . ."

The inference here is that it is reasonable to assume that permitted left-turns have the same arrival pattern as their adjacent through movements. Hence, in the absence of better knowledge about arrival patterns, it is suggested that the progression factors used for the eastbound and westbound left-turn lane groups be the same as those used for the adjacent through movement.

As a means of evaluating the impact of the revised approach, the delays estimated by it can be compared with those from the original analysis of Calculation 3. As can be observed from Table 3, the revised estimates of delay vary considerably from the original HCM estimates.

The most significant change can be observed for the northbound and southbound left-turn groups. The northbound delay has decreased by 69.5 percent whereas the southbound delay has increased by 38.1 percent. The reason for the decrease in northbound delay can be attributed to the additional permitted capacity calculated by Equation 2.

The increase in southbound delay is the result of the revised approach for calculating protected/permitted delay for individual phase components. The original approach reasoned that the total lane group delay could be estimated using the combined g/C ratio for the entire protected plus permitted phase (*1*, p. 9-56). However, this ratio will typically overestimate the true g/C ratio and result in unrealistically low uniform delay estimates for the left-turn lane group.

TABLE 3 DELAY COMPARISON

Lane Group	HCM Delay	Revised Delay	Change (%)
EB Left	36.0	34.6	-3.9
EB Thru	27.5	28.4	3.3
WB Left	36.0	34.0	-5.6
WB Thru	42.0	46.6	11.0
NB Left	71.4	20.1	-71.8
NB Thru	21.1	24.8	17.5
SB Left	54.6	72.7	33.2
SB Thru	7.4	6.1	-17.6
EB Approach	28.6	29.2	2.1
WB Approach	41.1	44.7	8.8
NB Approach	24.5	24.4	-0.4
SB Approach	14.7	16.4	11.6
Intersection	25.1	26.2	4.4

Differences in delay for other lane groups are not as great as for those of the northbound and southbound left-turn groups. These differences are small and can be attributed to slight changes in the analysis worksheet variables. For instance, the primary reason for the lower delay estimates for the eastbound and westbound left-turn lane groups is the different factor used to account for progression (i.e., 0.85 instead of 1.00).

CONCLUSIONS AND RECOMMENDATIONS

The main implication of this reanalysis is that there are inconsistencies between the original analysis of Calculation 3 and the HCM methodology. These are most likely misinterpretations of the HCM methodology that result from the general nature of the discussion related to protected/permitted left-turn phasing. In particular, the calculation of permitted capacity and unsaturated green time for exclusive, protected/permitted left-turn lane groups needs further clarification. Moreover, there is a need for clarification of the proper approach to use in estimating (a) total lost time, (b) amount of left-turn volume to assign to the permitted portion of protected/permitted left turns, and (c) delay for protected/permitted left-turn lane groups. Finally, it is recommended that Calculation 3 be amended to show both the initial and final worksheets thereby illustrating the iterative process involved in completing the capacity analysis worksheet.

REFERENCES

1. *Special Report 209: Highway Capacity Manual.* TRB, National Research Council, Washington, D.C., 1985.
2. J. H. Shortreed. Left-Turn Equivalencies for Opposed, Shared, Left-Turn Lanes. In *Transportation Research Record 971*, TRB, National Research Council, Washington, D.C., 1984, pp. 14–20.
3. D. B. Fambro, C. J. Messer, and D. A. Anderson. Estimation of Unprotected Left-Turn Capacity at Signalized Intersections. In *Transporiation Research Record 644*, TRB, National Research Council, Washington, D.C., 1977, pp. 113–119.

DISCUSSION

DANIEL F. BEAGAN
Stoner & Webster Engineering Corporation, P.O. Box 2325, Boston, Mass. 02107.

Bonneson and McCoy have made an important contribution toward clarifying the confusing aspects of the methodology for the analysis of signalized intersections as outlined in the *1985 Highway Capacity Manual (1).* Their recommendations for the total lost time, the progression factor for permitted left turns, and the general need for an iterative process in completing the worksheets should be incorporated directly in future updates of the manual. Their recommendations for opposing volume, the permitted portion of capacity, and for the use of effective green time in the supplemental worksheet for left-turn adjustment factors make valid points but need further discussion.

OPPOSING VOLUME

The authors propose quite reasonably that the opposing volume, V_o, be adjusted by the lane utilization factor, *LU*. It is this V_o that is then used in calculating the unsaturated green time, g_u. However, when the authors later calculate the capacity of the north and southbound left-turn permitted phases the unadjusted opposing volume is used. If this is not simply an error, the reason for using the unadjusted opposing volume at this point should be given.

EFFECTIVE PERMITTED GREEN TIME

In preparing the supplemental left-turn adjustment factor worksheet for the northbound permitted phase, the authors used the value of 80.7 sec, which is the total opposing effective green time including both the protected and permitted left-turn phases. This is necessary to obtain the proper unsaturated green time value because the saturated green includes the opposing southbound volume, which moves on the protected phase. Although the authors did not fill out the complete worksheet, the equations for g_g and f_m, which both involve g, should use the northbound *permitted* left-turn green time of 70 sec, not 80.7. This argues for the need to include a new input variable for the worksheet, g_o, which is equal to the total green time associated with the opposing volume. It would be used in place of g in calculating the unsaturated green time.

CAPACITY OF PERMITTED LEFT-TURN PHASE PORTION

The authors correctly conclude that the term g used to calculate the capacity of the permitted portion of the protected/permitted left turn is actually the unsaturated green time, g_u. This maintains consistency with the left-turn saturation flow adjustment factor for permitted left turns. The authors, however, continue to follow the manual in choosing the maximum of the unsaturated portion or the change interval capacity for the permitted capacity of the protected/permitted left turn. For complete consistency with the supplemental left-turn adjustment factor worksheet, the sum, not the maximum, of these capacities should be used, and that sum should be multiplied by all of the adjustment factors from the saturation flow adjustment worksheet contained in Columns 5 through 11. Failure to do this will result in a different capacity for the permitted left-turn phase, depending on whether it is handled alone or as part of a protected/permitted left turn.

DELAY FOR PROTECTED/PERMITTED LEFT TURNS

The authors display the results of a uniform delay analysis based on calculating the area under the queue-departure diagram as shown in Figure 8. Because this method was actually used in the original formulation of the uniform delay equation, d_1, this is the correct procedure. Those using the method should be cautioned, however, that, to be consistent with the manual, the area must be reduced by 33 percent as the authors have done to account for the conversion from total delay to stopped delay. Further, although not noted in Figure 8, the downward sloping lines are all the difference between the saturation flow rate and the arrival rate.

Finally, in Figure 8 the end of the northbound queue occurs at 83.0 not 100.1 sec as shown, although the average uniform delay is correct. The southbound queue after 32.3 sec is 3.4 not 3.7 vehicles, and the southbound queue is reduced to zero at 41.5 sec, not 44.8 as shown. For the southbound lane group, the area needs to be recalculated. The total average uniform delay for the southbound left-turn lane groups should be 28.5 not 29.8 sec shown in Figure 7.

DISPLAY OF RESULTS

In Figure 7, the level of service worksheet, the authors display the saturation flow rates and green ratios differently for the north-south and the east-west left-turn groups. For the east-west lane groups, they show the total green time ratio and a saturation flow rate accounting for the total green time. For the north-south left-turn lane groups they show the unsaturated green time ratios and a saturation flow rate accounting for the unsaturated green time ratio. Although the delay results obtained will be the same regardless of whether the total or unsaturated green ratios are used if the proper associated saturation flow rates are also used, the different presentations may introduce unnecessary confusion in the review of completed worksheets.

CONCLUSION

The authors have made an important step in correcting the deficiencies of the signalized analysis methodology of the *1985 Highway Capacity Manual.* The methodology for constructing the queue-departure diagram needs to be more fully described in the future. The capacity calculation of the permitted portion of a potential/permitted left turn also needs further discussion.

AUTHORS' CLOSURE

As Beagan notes in his discussion, we have identified several areas in need of clarification or revision to the signalized intersection analysis methodology in the *1985 Highway Capacity Manual* (HCM). Many of our recommendations were for further clarification of areas that were vague in their application toward protected/permitted left-turn phasing. These include the calculation of unsaturated green time and capacity of the permitted phase portion, the progression factor adjustment, assignment of left-turn volume to each phase portion, and the general iterative nature of the capacity analysis. On the other hand, there are some areas that would appear to need revision. These include the calculation of lost time and delay with regard to protected/permitted movements.

Beagan appears to agree with many of our findings while taking issue with others. In general, his discussion highlights several points that perhaps were not discussed as exhaustively in our paper as they could have been. We are hopeful that his comments will provide any further clarification needed in those areas.

With regard to Beagan's discussion, we would like to offer some additional comment. In particular, he suggests that further explanation is required about the use of the unadjusted opposing volume (i.e., not adjusted for lane utilization) in calculating the capacity of the permitted phase portions. The omission of this adjustment was intentional and reflects the authors' understanding of the derivation of the equation used to calculate permitted left-turn saturation flow rates (i.e., Equation 2).

It is believed that this equation is a linear approximation of the negative exponential function derived originally by Major and Buckley to describe the interaction of two traffic streams at a priority intersection *(1)*. This particular equation does not explicitly account for the number of opposing lanes although there are several equations that do *(2)*. More important, however, none of these equations uses a lane utilization factor as a means of addressing the number of opposing lanes.

It should also be noted that a lane utilization adjustment has been traditionally used to account for unequal lane use by a queue of vehicles on an intersection approach. This adjustment is intended to account for the green time required to clear the longest standing queue. It is our opinion that lane utilization adjustments are inappropriate for determining the permitted left-turn saturation flow rate of vehicles filtering through a randomly arriving stream. It should be noted that the approach used is consistent with that of others *(3)*.

Beagan also suggests that the combined capacities of the end-of-phase "sneakers" and the unsaturated phase portion be used for the permitted left-turn capacity instead of simply using the larger of the two. We would agree that in many cases both of these components combine to serve existing left-turn demand and should be analyzed as such. However, the design of a signal timing plan that does not adequately serve the total left-turn demand, without relying on sneakers, should not be recommended. Any timing plan that is designed to take advantage of sneaker activity encourages improper use of the change interval, increases the number of vehicle conflicts, and compromises the safety of all motorists within the intersection.

The approach developed in our paper is consistent with the HCM's methodology and is in recognition of the aforementioned concerns. Using this approach, enough green time would be provided the left-turn phase to serve all left-turn vehicles except those that clear during the unsaturated phase portion. This approach would minimize the amount of "sneaker" activity. However, in special cases where "sneakers" provide the greater permitted capacity (i.e., when $g_u = O$ or $V_o > 1,399$), there will undoubtedly be some vehicles moving at the end of the phase. In this situation, it would be advisable to use protected-only instead of protected/permitted left-turn phasing.

Beagan also comments on the calculation of the uniform and random delay components. In fact, his comments have brought to light the need for some minor changes to the uniform delay estimates and Figures 7 and 8. These revisions were made for the final version of our paper.

REFERENCES

1. N. G. Major and D. J. Buckley. Entry to a Traffic Stream. *Proc., Australian Road Research Board,* Nunawading, 1962.
2. P. G. Michalopoulos, J. O'Connor, and S. M. Nova. Estimation of Left-Turn Saturation Flows. In *Transportation Research Record 667,* TRB, National Research Council, Washington, D.C., 1978, pp. 35–41.
3. D. B. Fambro, C. J. Messer, and D. A. Anderson. Estimation of Unprotected Left-Turn Capacity at Signalized Intersections. In *Transportation Research Record 644,* TRB, National Research Council, Washington, D.C., 1977, pp. 113–119.

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

Utilization and Timing of Signal Change Interval

FENG-BOR LIN, DONALD COOKE, AND SANGARANATHAN VIJAYAKUMAR

The problem of timing the signal change interval has received increased attention in recent years. Much of this attention is focused on two issues: whether a constant yellow interval should be used and whether the timing equations suggested by the Institute of Transportation Engineers (ITE) can realistically reflect driver needs for the change interval. These two issues are examined on the basis of observed driver behavior. The 95th percentile yellow interval requirements are found to vary from 3 to 5 sec. Such requirements do not have a positive linear correlation with the approach speed. The 85th and 95th total change interval requirements have strong linear correlations with vehicle clearance time. The ITE's timing equations should be replaced by simpler ones that can better explain driver behavior.

A signal change interval is a short time period in a traffic signal cycle between conflicting green intervals. A yellow signal indication is displayed in this interval, which is often followed by an all-red signal indication. There are two major problems in timing the signal change interval for the vehicles on an intersection approach. One is to determine the total change interval requirement, and the other is to divide this total requirement into the yellow interval and the all-red interval requirement.

In general, the total-change interval requirement refers to the length of a change interval needed for a safe transfer of the right-of-way. The yellow interval requirement represents the length of a yellow interval that is needed to allow a reasonable driver to take proper action before a red signal indication is exhibited. The all-red interval requirement is the additional time following a yellow interval that is needed to clear vehicles from the intersection before a green signal indication is displayed for the vehicles on other approaches.

Current practices in determining these various requirements associated with the signal change interval vary among traffic engineering agencies. Nevertheless, the following equation suggested by the Institute of Transportation Engineers (ITE) *(1)* has been adopted by many agencies for determining the change interval requirement:

$$T = t + V/(2a) + (W + L)/V \qquad (1)$$

where

T = change interval requirement, in sec;
t = driver reaction time, in sec;
V = vehicle approach speed, in ft/sec;
a = vehicle deceleration rate, in ft/sec^2;
W = intersection width, in ft; and
L = vehicle length, in ft.

The sum of the first two terms on the right side of this equation has also been used by a number of agencies in determining the yellow interval requirement.

The use of Equation 1 requires the selection of representative values for the driver reaction time, vehicle deceleration rate, and vehicle length. Reported values of the mean and the 85th percentile reaction times and deceleration rates vary significantly from one intersection to another *(2)*. ITE suggests a value of 1 sec for the reaction time and 10 ft/sec^2 for the deceleration rate. The value for the vehicle length is commonly assumed to be 20 ft.

Equation 1 has been expanded in several studies *(3, 4)*. In May 1985, ITE *(5)* also extended this equation and proposed a recommended practice in timing the change interval. This recommended practice determines the yellow interval according to

$$Y = t + V/2/(a \pm 0.322G) \qquad (2)$$

where t, V, and a are as defined for Equation 1, and Y = yellow interval requirement, in sec; and G = the grade of approach lane, in percent. ITE recommends that the 85th percentile approach speed always be used in Equation 2.

In addition, the ITE's proposed recommended practice allows the use of $(W + L)/V$, or P/V, or $(P + L)/V$ to determine the length of the required all-red interval. The notations used in these terms are to be interpreted as follows:

W = width of the intersection, measured from the near-side stop line to the far-side edge of the conflicting traffic lane along the actual vehicle path, in ft;
P = width of intersection, measured from the near-side stop line to the far side of the farthest conflicting pedestrian crosswalk along the actual vehicle path, in ft;
L = length of vehicle, recommended as 20 ft; and
V = speed of the vehicle through the intersection, in ft/sec.

According to ITE, $(W + L)/V$ is to be used if there is no pedestrian traffic present; the longer of $(W + L)/V$ and P/V should be used if there is the probability of pedestrian crossings, and $(P + L)/V$ should be used if there is significant

Civil and Environmental Engineering Department, Clarkson University, Potsdam, N.Y. 13676.

pedestrian traffic or the crosswalk is protected by pedestrian signals.

To determine the entire change interval, the ITE's proposed recommended practice requires that the sum of the yellow interval and the all-red interval be calculated twice—once with the 15th percentile speed and again with the 85th percentile speed. If the 15th percentile speed produces a longer interval, the all-red interval calculated at the 85th percentile speed is to be increased by the difference.

Several recent studies have raised doubt about the wisdom of using Equation 1 or its expanded forms in determining the change interval and of using Equation 2, or the sum of t and $V/(2a)$, for determining the yellow interval. For example, Chang et al. have found that the behavior of the drivers who entered the intersection after the yellow onset did not change significantly with the approach speed *(6)*. Their study showed that, over a speed range of 25 to 55 mph, 85 percent of the vehicles entering the intersection after the yellow onset took less than approximately 3.5 to 3.8 sec to reach the stop line. And, over the same speed range, 95 percent of the entering vehicles took less than about 4.2 to 4.6 sec to reach the stop line after the yellow onset. These findings prompted the three investigators to suggest that the use of a constant yellow interval of 4.5 sec may be warranted. A study by Wortman and Fox further reinforces the notion that the needs for the yellow interval is independent of the approach speed *(7)*.

Regarding the length of the change interval, a study by Lin has shown that the change interval requirement can be better estimated as a linear function of the time required for the vehicles to clear the intersection *(2)*. However, Lin's study was based on a rather limited data base. Subsequent to this study, additional data were collected in order to provide a better understanding of how the change interval should be designed.

The objective of this paper is to use the available data to discuss the utilization and timing of both the yellow interval and the change interval as a whole.

YELLOW INTERVAL REQUIREMENTS

Faced with a yellow signal indication, a driver will either decide to stop or proceed through the intersection. A yellow interval should be long enough to allow a proper choice by a driver under such a circumstance. Whether or not a yellow interval is adequate can be evaluated in terms of the percent of vehicles entering the intersection after the termination of the yellow interval *(5)*. A shorter yellow interval will likely force a larger percent of vehicles to enter the intersection on a red signal indication, or force more drivers to take potentially dangerous actions. The yellow interval requirement is defined in this study as the length of a yellow interval that will allow a specified percent of signal change intervals to be free of vehicles entering the intersection on a red signal indication.

To examine the nature of this requirement, data related to six straight-through movements and two left-turn movements at a total of five intersections were collected for analysis. Each of the subject movements represents the vehicular flows in one or more traffic lanes. All the five subject intersections were located in the state of New York. Three were on Central Avenue in Albany, one was on Almond Street in Syracuse, and the remaining one was on Market Street in Potsdam. The two left-turn movements had their own separate signal phases. As can be observed in Table 1, the clearance widths for the eight movements varied from 77 to 135 ft. Pedestrian interferences were negligible at the time of the data collection. Therefore, for each of the straight-through movements, the clearance width was measured from the stop line of the approach lane to the farthest potential conflicting point on the far side of the intersection. Similarly, the clearance widths for the two left-turn movements were measured as the length of a representative turning path from the stop line to the farthest potential conflicting point downstream.

The approach speeds given in Table 1 were based on those vehicles approaching the intersection near the end of the green interval. They were measured with stopwatches as the travel times over a distance of 100 to 150 ft. The lowest mean approach speed was 21.9 mph for Movement 8 and the highest was 32.5 mph for Movement 4. The grades for all the movements were gentle.

On average, each of the subject movements was observed for about 2-1/2 hr. The number of change intervals encountered in such an observation period ranged from 68 for Movement 8 to 255 for Movement 6. Not every one of these change intervals was utilized by vehicles either to enter or to clear the intersection after the yellow onset. For each utilized change interval, the elapsed time from the yellow onset to the moment the last entering vehicle reached the stop line was measured with a stopwatch. Such an elapsed time represents the length of the yellow interval that is needed for a change interval to be free of vehicles entering on a red signal indication. The resulting

TABLE 1 CHARACTERISTICS OF MOVEMENTS EXAMINED FOR YELLOW INTERVAL REQUIREMENTS

Movement	Movement Type	Existing Yellow sec	Clearance Width w, ft	Approach Speed, V mph			Grade %
				15th	Mean	85th	
1	Straight	3.9	77	25.6	30.0	41.9	+3.0
2	Straight	3.9	135	26.2	29.8	36.0	+1.7
3	Straight	4.0	105	22.1	27.6	38.0	+0.9
4	Straight	3.5	96	26.6	32.5	37.8	+0.7
5	Straight	3.5	92	22.6	26.6	30.9	+0.8
6	Straight	3.1	93	27.0	31.8	38.2	+0.9
7	Left	3.0	115	21.9	26.3	31.7	+0.9
8	Left	3.9	105	18.2	21.9	24.6	-0.6

measurements for each subject movement were used to construct a cumulative distribution of the yellow interval requirement. Figure 1 shows the cumulative distributions of the yellow interval requirements for the eight subject movements. Based on these distributions, a yellow interval can be chosen that will allow a reasonably high percent (e.g., 85 or 95 percent) of the signal cycles to be free of vehicles entering on a red signal. The 85th and 95th percentile yellow interval requirements for the eight subject movements are given in Table 2 along with related statistics.

The 95th percentile yellow interval requirements varied from about 3 to 5 sec, and the 85th percentile requirements were between 2.2 and 4.2 sec. These variations cannot be explained by the differences in the approach speeds of the various movements. Figure 1 shows that it can be quite erroneous to assume that the yellow interval requirements has a positive linear correlation with the approach speed.

Among the eight movements examined, Movement 4 and Movement 8 (Table 1) had the largest difference in approach speed. Thus, if the approach speed governs the yellow interval requirement in a manner as indicated by Equation 2, the cumulative distributions of the yellow interval requirements of these two movements should have exhibited the largest difference. To the contrary, Figure 1 shows that they were nearly identical. On the other hand, Movement 4 and Movement 6 had virtually the same approach speed. Yet, their yellow interval requirements displayed a large difference. Movement 4 and Movement 5 did show an increase in the yellow interval requirement as the approach speed increased. But, the yellow interval requirements of Movement 7 and Movement 8 ex-

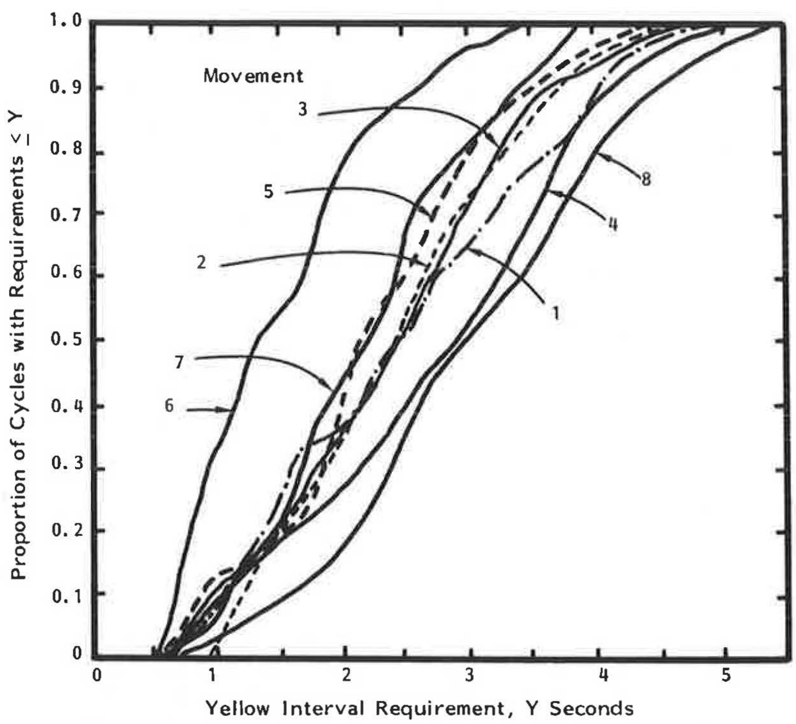


FIGURE 1 Cumulative frequency distributions of yellow interval requirements.

TABLE 2 YELLOW INTERVAL REQUIREMENTS

Movement	Requirement in Seconds 85th	Requirement in Seconds 95th	Mean Approach Speed mph	Change Interval Number Utilized	Change Interval Percent Utilized
1	3.9	4.2	30.0	55	49
2	3.5	4.0	29.8	55	74
3	3.4	4.2	27.6	60	70
4	3.9	4.5	32.5	101	76
5	3.2	3.9	26.6	55	59
6	2.2	3.0	31.8	74	29
7	3.2	3.7	26.3	57	46
8	4.2	5.0	21.9	54	79

hibited a relationship contradictory to that implied in Equation 2.

Thus, a question can be raised as to what really accounted for the variations in the cumulative frequency distributions shown in Figure 1. With each additional hour of field observations made by the authors, it became increasingly clear that the supply of vehicles that were in a position to enter the intersection within 5 sec after the yellow onset was a major source of such variations. At one extreme, Movement 8 (left turns from Wolf Road onto Central Avenue in Albany) had frequent carryovers of long queues from one cycle to the next because of the inability of a rather short green interval to discharge all of the queueing vehicles in a cycle. As a result, the vehicles would often continue entering the intersection long after the yellow interval (3.9 sec) expired.

Similarly, Movement 4 (straight-through flows on Almond Street at Harrison Street in Syracuse) provided a high level of vehicle supply at the time of the data collection. This movement occupied three straight-through lanes and another lane shared by straight-through and right-turn vehicles. During the evening peak hours in which most of the observations were made, a large number of vehicles were frequently within short travel times from the stop line at the yellow onset, and long queues often began to develop immediately after the change interval expired. Consequently, the yellow interval requirement of this movement differed very little from that of Movement 8.

At the other extreme, Movement 6 (a straight-through flow on Market Street at Sandstone Road in Potsdam) had a low flow rate of about 300 vph and was regulated by a traffic-actuated signal. The level of vehicle supply at the yellow onset was low because the vehicles were usually more than 4 sec away from the intersection when the green interval expired. Over a 4-hr observation period, the longest recorded yellow interval requirement for a cycle was 3.4 sec, and the 95th percentile yellow interval requirement was only 3 sec. Movement 7 (left turns from Central Avenue onto Everlet Avenue in Albany) had a similar characteristic. The yellow interval for this movement often began when there were no vehicles within a travel time of less than 4 sec from the stop line. This phenomenon was created by the signal controls for Movement 7 and for the movements at the upstream intersection. The resulting 95th percentile yellow interval requirement of 3.7 sec was significantly lower than that of Movements 4 and 8.

Although it is evident that the level of vehicle supply at the yellow onset is a governing factor of the yellow interval requirement, there are currently no quantitative methods for defining such a causal relationship. The percent of change intervals utilized by vehicles to enter the intersection may be a potential measure of the level of vehicle supply. Figure 2 shows that such a measure has an apparent correlation with both the 95th and 85th percentile yellow interval requirements of the eight movements.

Let F be the proportion of change intervals utilized by vehicles to enter the intersection after the yellow onset. Then, the 95th percentile yellow interval requirements of the eight movements given in Table 2 can be related to F according to

$$Y = 2.36 + 2.83F \quad (3)$$

where Y represents the specified percentile yellow interval requirement. This equation has an R^2 value of 0.73 and a standard error of estimate of 0.33 sec. The corresponding equation for the 85th percentile yellow interval requirement is

$$Y = 1.81 + 2.70F \quad (4)$$

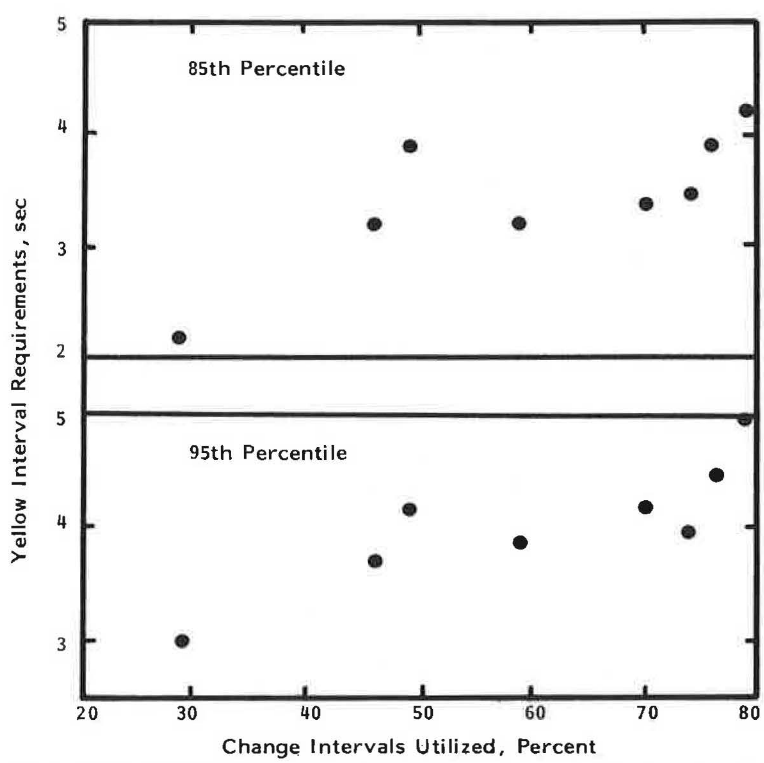


FIGURE 2 Variation of yellow interval requirement with the rate of change interval utilization.

The R^2 value of this equation is 0.60 and the standard error of estimate is 0.42 sec.

CHANGE INTERVAL REQUIREMENTS

The signal change interval may comprise only a yellow interval or include a yellow interval and an all-red interval. The current Uniform Vehicle Code *(8)* allows vehicles to enter the intersection during the yellow interval and to clear the intersection after the red interval begins. This "permissive rule" has a greater need for the all-red interval in comparison with the "restrictive rule," which requires vehicles to clear the intersection by the end of the yellow interval.

Regardless of which rule drivers should follow, the change interval requirement can be defined as the length of a change interval that is needed to allow all vehicles to clear the intersection in a specified percent (e.g., 85 percent) of signal cycles. In order to analyze the nature of this requirement, data related to the interactions between the change intervals and the vehicles of 22 movements were collected. These 22 movements were associated with 15 intersections in 5 urban areas in the state of New York. Five of the intersections were located in Syracuse, four in Albany, one in Rochester, three in Potsdam, and two in Canton. Four of the 22 movements were the same as Movement 1 through Movement 4 described previously in Table 1.

As can be observed in Table 3, all but one of the movements had clearance widths between 74 and 135 ft. The grades of the approach lanes were within ±4 percent. The mean approach speeds varied from 21.9 to 43.9 mph, and the mean turning speeds of the left-turn movements were about 20 mph. At the time of the data collection, none of the 22 movements had vehicles blocking the intersections because of congestion downstream of the stop lines. The "permissive rule" was in place at all the intersections for the use of the change interval. Pedestrian interferences with the vehicular movements at all the intersections were negligible at the time of the data collection. Therefore, the clearance widths were also measured according to the definitions described previously.

The traffic conditions in an approach lane of a signalized intersection may vary substantially within a signal cycle because of the formation and dissipation of queues. Consequently, the speeds of those vehicles that may interact with the change interval may be significantly affected by such changing conditions. For this reason, the determination of the vehicle speeds took into consideration only those vehicles approaching or crossing the intersection near the end of the green interval or immediately after the yellow onset. Stopwatches were used to measure the travel times of such vehicles over a distance of 100

TABLE 3 CHANGE INTERVAL REQUIREMENTS

Movement	Clearance Width ft	Grade %	Approach Speed, mph			C.I. Requirement, sec		% of C.I. Utilized
			15th	Mean	85th	85th	95th	
1	89	-1.0	25.7	28.9	32.3	6.3	7.0	51
2	89	+4.0	23.8	27.2	35.2	5.6	6.4	60
3	117	-0.5	27.5	31.1	35.8	5.0	6.0	60
4	74	-0.9	24.3	27.9	32.0	4.8	5.5	36
5	107	-0.3	28.4	32.2	36.5	5.3	6.0	15
6	106	-3.9	27.6	30.1	31.7	5.5	6.0	34
7	90	+0.7	38.6	43.9	49.2	5.5	5.8	47
8	96	+0.2	28.1	33.0	37.6	5.0	5.5	61
9	195	+1.0	24.2	30.6	35.8	7.4	8.3	23
10	74	+0.4	27.3	30.8	35.1	5.1	5.7	55
11	130	+0.9	27.7	32.5	38.1	5.8	6.1	40
12	77	+3.0	25.6	30.0	41.9	5.6	6.3	46
13	135	+1.7	26.2	29.8	36.0	6.8	8.3	79
14	76	-0.1	23.5	28.2	35.6	5.7	6.6	60
15	105	+0.9	22.1	27.6	38.0	7.2	8.1	67
16	110	-3.5	20.8	24.5	28.9	6.8	8.0	44
17	96	+0.7	26.6	32.5	37.8	5.4	6.0	66
18	130	+0.6	25.9	30.5	35.7	6.2	7.1	39
19	105	+0.8	17.9 (14.9)	21.9 (18.0)	25.2 (22.7)	8.8	9.3	79
20	115	+0.9	22.6 (17.4)	26.3 (20.0)	31.1 (24.4)	7.4	8.0	26
21	96	-0.5	24.3 (17.9)	28.3 (20.2)	33.6 (23.2)	7.4	7.8	44
22	101	+0.8	22.5 (16.2)	26.8 (18.8)	30.8 (21.1)	8.4	9.4	69

Note: Values in parentheses are turning speeds to 150 ft. For the straight-through movements, the approach speeds determined in this manner were equivalent to the speeds at which the vehicles cleared the intersection. For the left-turn movements, the approach speeds were not a good approximation of the clearance speeds. Therefore, both approach speeds and clearance speeds were determined for the left-turn movements.

The major task of the data collection was to use stopwatches to measure, on a cycle-to-cycle basis, the elapsed time from the yellow onset to the moment the last entering vehicle cleared the intersection. Only those vehicles entering the intersection after the yellow onset were included in the data collection. This task was performed for an average of about 2 hr for each subject movement. The resulting data were used to construct a cumulative frequency distribution of the change interval requirement for each subject movement. The 85th and 95th percentile change interval requirements determined from such distributions are summarized in Table 3 along with other relevant statistics.

The data given in Table 3 can be used to examine alternative models for estimating the change interval requirement. One such model suggested by Lin *(2)* can be written as

$$T = A + B(W + L)/V \tag{5}$$

where

T = specified percentile requirement of the change interval, in sec;
A, B = coefficients to be calibrated;
W = clearance width, in ft;
L = representative vehicle length, 20 ft; and
V = mean clearance speed, in ft/sec.

This model implicitly assumes that the relationship between the change interval requirement and the average clearance time $(W + L)/V$ is linear. Figure 3, which is based on the 95th percentile change interval requirements given in Table 3, confirms the existence of a strong linear relationship between T and $(W + L)/V$.

A least-square regression based on the 85th percentile change interval requirement given in Table 3 results in the following equation:

$$T = 2.84 + 1.09(W + L)/V \tag{6}$$

This equation has an R^2 value of 0.75 and a standard error of estimate of 0.58 sec. When the 95th percentile requirements are used for the regression, the resulting equation is

$$T = 3.33 + 1.17(W + L)/V \tag{7}$$

The R^2 value of this equation is 0.74 and the standard error of estimate is 0.64 sec.

No attempt was made to analyze the confidence intervals of these regression equations and the variances of the regression coefficients. Such an analysis requires the assumption that the 85th and 95th percentile change interval requirements are distributed normally. The existing data do not support such an assumption.

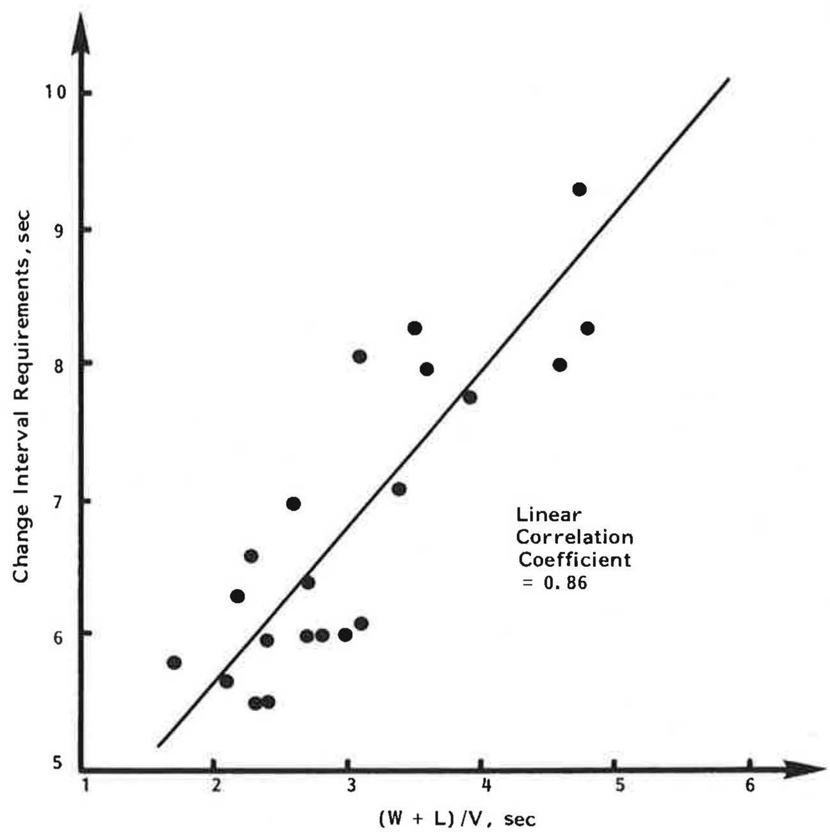


FIGURE 3 Variation of the 95th percentile change interval requirement with clearance time.

For the subject movements, the last term of Equation 7 is virtually the same as the clearance time determined from the 15th percentile clearance speed. Similarly, the last term of Equation 6 has values that are on the average only 0.2 sec shorter than the clearance times determined from the 15th percentile clearance speeds of the various movements. Therefore, if the last terms of these two equations are used to determine the all-red interval requirement, they would satisfy the clearance needs of about 85 percent of the entering vehicles.

Equations 6 and 7 are capable of explaining about 74 percent of the variations in the 85th and the 95th percentile change interval requirements. The unaccounted-for variations can be attributed to differences in vehicle speeds, vehicle lengths, vehicle supply patterns at the yellow onset, and so forth. Because the yellow interval requirement is linearly correlated to some extent with the proportion F of change intervals utilized, Equation 5 can be improved by adding onto it another term as follows:

$$T = A + CF + B(W + L)/V \tag{8}$$

The resulting regression equation based on the 95th percentile requirements is

$$T = 2.24 + 2.15F + 1.18(W + L)/V \tag{9}$$

with an R^2 value of 0.84 and a standard error of estimate of 0.52 sec. The corresponding equation for the 85th percentile requirements is

$$T = 2.0 + 1.7F + 1.10(W + L)/V \tag{10}$$

This equation has an R^2 value of 0.82 and a standard error of estimate of 0.51 sec.

Although Equations 9 and 10 are more powerful than Equations 6 and 7 in explaining the length requirements of the change interval, the improvements do not appear to be large enough to warrant the use of either Equation 9 or Equation 10. In fact, the inclusion of F in these equations would make the equations difficult and expensive to use because data for F would have to be collected at each intersection. For the same reason, Equations 3 and 4 presented earlier would have little use for timing applications.

An alternative to Equations 5 and 8 for estimating the change interval requirements is ITE's proposed recommended practice. For movements that have little pedestrian interference, this practice implies a model form of

$$T = t + V/2/(a \pm 0.322G) + (W + L)/V \tag{11}$$

As described previously, the use of Equation 11 requires two calculations: once with the 15th percentile speed and once with the 85th percentile speed. However, ITE is vague about how such calculations are to be performed for protected left turns. The last two terms in Equation 11 are a function of vehicle speed. If both terms are determined from the same percentile approach speed for left-turn movements, it can be shown that the resulting T values and the 95th percentile change interval requirements given in Table 3 have a linear correlation coefficient of 0.57. If the approach speed is used for the second term on the right side of Equation 11 and the turning speed is used for the last term, the resulting linear correlation coefficient would become 0.79 (Figure 4). This level of correlation with the observed 95th percentile requirements is respectable, but it is still below that which can be achieved by a much simpler model such as Equation 7 (Figure 3). Therefore, there is no reason to adopt ITE's proposed recommended practice unless it has a superior theoretical basis for explaining driver behavior.

The soundness of the theoretical basis for Equation 11 can be addressed by rewriting this equation as

$$Z = T - (W + L)/V = t + V/2/(a \pm 0.322G) \tag{12}$$

The term Z in this equation represents the yellow interval requirement. If Equation 11 is a valid representation of the behavior of drivers in their use of the change interval, the values of Z determined as $T - (W + L)/V$ from field observations should be strongly and positively correlated with the approach speed.

Figure 5 shows that such a linear correlation does not exist between Z and the approach speed as far as the 22 subject movements are concerned. In this figure, the values of Z are determined as $T - (W + L)/V$ from Table 3 based on the mean clearance speeds and the 95th percentile change interval requirements. A least-square regression of these Z values on the mean approach speeds (in mph) results in the following equation:

$$Z = 5.32 - 0.049V \tag{13}$$

The R^2 value of this equation is 0.10 and the standard error of estimate is 0.62 sec.

As can be observed from Figure 5, there is only one Z value for mean approach speeds exceeding 34 mph. This Z value is weighted more heavily in Equation 13 than the other values. If this value is deleted, the resulting regression equation becomes

$$Z = 7.85 - 0.14V \tag{14}$$

for mean approach speeds ranging from 22 to 33 mph. Equation 14 has an R^2 value of 0.34 and a standard error of estimate of 0.54 sec.

The regression coefficients of both Equation 13 and Equation 14 are quite different from the values recommended by the ITE. According to ITE, the constant terms in Equations 12 and 13 should have been 1.0 sec and, for the subject movements, the coefficient of V should have been in the range of 0.044 to 0.057. Therefore, even if Z is really a linear function of the approach speed, the sum of t and $V/2/(a \pm 0.322G)$ is a poor representation of driver behavior. The negative signs of the coefficients of V in Equations 13 and 14 further indicate that an increase in the approach speed tends to cause a reduction instead of an increase in the yellow interval requirement. It is uncertain, however, whether this negative correlation between Z and V really exists because the R^2 values of Equations 13 and 14 are rather small and data are lacking for mean approach speeds exceeding 32 mph and for speeds below 26 mph. Overall, it is evident that the causal relationship between Z and V is very weak and; thus, it is meaningless to treat the yellow interval requirement as a function of the approach speed.

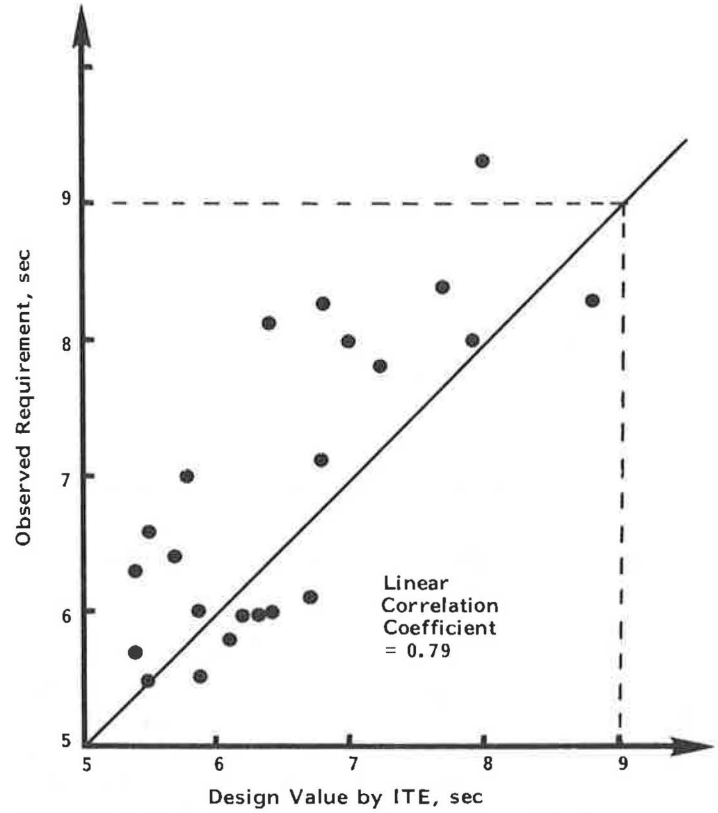


FIGURE 4 Correlation between the 95th percentile change interval requirements and values determined from the ITE's proposed recommended practice.

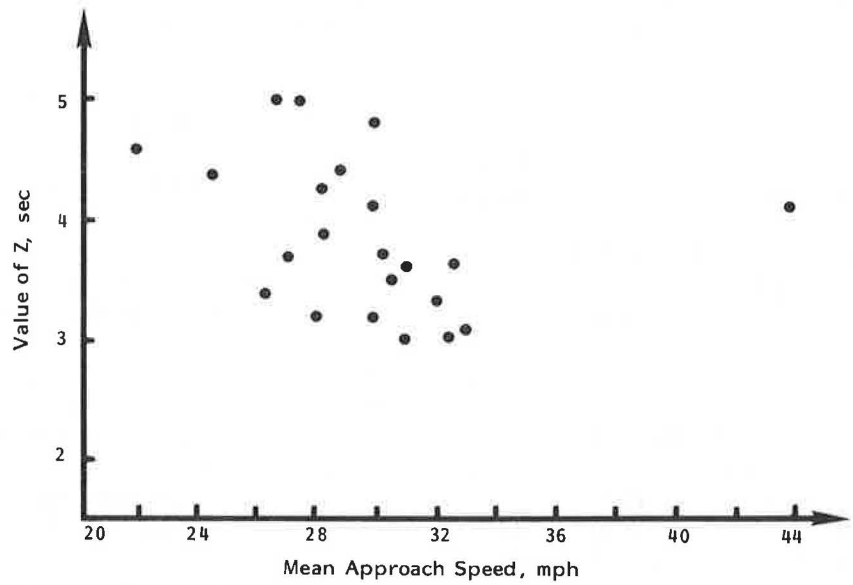


FIGURE 5 Variation of Z with mean approach speed.

TIMING DESIGN APPLICATIONS

The cumulative frequency distributions of the yellow interval requirement shown in Figure 1 are bounded by the distributions of Movements 6 and 8, which had vehicle supply patterns of opposite extremes at the yellow onset. The corresponding 95th percentile yellow interval requirements are between 3 and 5 sec and the 85th percentile requirements are between 2.2 and 4.2 sec. The *Z* values shown in Figure 5, which approximate the 95th percentile yellow interval requirements of 22 movements, also lie between 3 and 5 sec. The same *Z* values plotted in Figure 6 against the percentage of change intervals utilized further show that the *Z* values remain above 3 sec even when the rate of change interval utilization drops to as low as 13 percent. Therefore, a reasonable range of the design values for

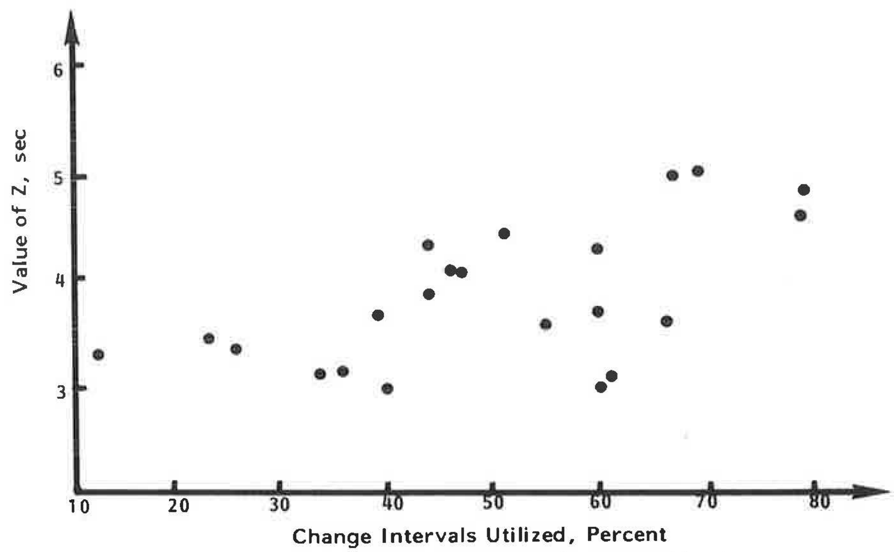


FIGURE 6 Variation of Z with the rate of change interval utilization.

the yellow interval is 3 to 5 sec. Yellow intervals shorter than 3 sec are not recommended because they may cause some drivers to apply excessively high decelerations in order to avoid entering the intersection on red.

The variations in the change interval requirements cannot be accounted for by the differences in the approach speeds. Relating such variations to another traffic or signal variable is likely to make the resulting timing method difficult to apply. Therefore, it is preferred that simple guidelines be established in the future for the choice of the yellow interval.

Meanwhile, the yellow interval may be determined according to

$$Y = 4.0 + C_1 \tag{15}$$

in order to satisfy the 95th percentile requirements. In this equation, C_1 is a correction factor with a value between -1.0 and $+1.0$ sec.

A value between $+0.5$ and $+1.0$ sec may be chosen for C_1 if one of the following vehicle supply patterns exists: (a) frequent carryovers of long queues from one cycle to the next; (b) a rapid build-up of long queues immediately after the change interval expires; and (c) the percent of change intervals utilized exceeding 70 percent. On the other hand, a reasonable choice of C_1 would be between -1.0 and -0.5 sec if (a) the movements of concern have low flow rates and are regulated by traffic-actuated controls; or (b) the vehicle supply to the intersection at the yellow onset is frequently cut off due to cyclic flow patterns created by signal coordination; or (c) the rate of change interval utilization is less than 30 percent. For movements with vehicle supply patterns in between the two extremes, C_1 may be set to 0 sec.

To reflect the actual requirements of individual movements, the lengths of the change interval determined from either Equation 6 or Equation 7 may be adjusted upward or downward. For the 95th percentile requirements, the adjustment may take the form of

$$T = 3.33 + 1.17\,(W + L)/V + C_2 \tag{16}$$

where C_2 is a correction factor.

For the 22 subject movements given in Table 3, the values of C_2 range from about -1.0 to $+1.1$ sec. These values correspond to the deviations of the measured 95th percentile requirements from the regression line shown in Figure 3. Again, a reasonable choice of C_2 is between $+0.5$ to $+1.0$ sec for movements with high levels of vehicle supply at the yellow onset (e.g., Movements 1, 7, 13, 14, 15, 16, 19, and 22). In contrast, C_2 would most likely assume a value between -1.0 and -0.5 sec for movements with very low levels of vehicle supply (e.g., Movements 3, 5, 8, 9, 11, and 20).

For timing applications, C_1 and C_2 can be considered to be the same. Therefore, given a yellow interval and a change interval as determined from Equations 15 and 16, the all-red interval R can be calculated as

$$\begin{aligned} R &= 3.33 + 1.17\,(W + L)/V + C_2 - 4 - C_1 \\ &= 1.17\,(W + L)/V - 0.67 \end{aligned} \tag{17}$$

It should be noted that the all-red interval requirements determined from Equation 17 do not take into consideration a safety margin that may be provided by the cross traffic. This safety margin is the amount of time required for the first vehicle in the cross traffic to reach the conflicting point after the change interval expires. If queueing vehicles are present on the cross street after the change interval expires, the safety margin would equal the time needed for the driver in the first queueing vehicle to accelerate the vehicle to the conflicting point. If no queueing vehicles are present, it will also take the first vehicle arriving on the cross street some time to reach the conflicting point.

If this vehicle approaches the intersection at a speed V_o and has an intended deceleration b, then the minimum safety margin provided by this vehicle can be approximated by $V_o/(2b)$. The deceleration rate b can be as high as 18 ft/sec^2 *(6)*. Using this rate and measuring V_o in ft/sec gives a safety margin of $V_o/36$ sec. This safety margin can be determined by choosing an appropriate value (e.g., the 15th percentile approach speed) for V_o. The safety margin allowed for the timing design could be taken as the lesser of $V_o/36$ and the minimum (or near minimum) time for the first queueing vehicle in the cross traffic to reach the conflicting point.

CONCLUSIONS

The yellow interval requirement correlates poorly with the approach speed. This requirement appears to be governed by the vehicle supply pattern at the yellow onset. When frequent carryovers of long queues from one cycle to the next exist, the 95th percentile yellow interval requirement can reach 5 sec. This requirement can be reduced to about 3 sec if vehicular movements with low flow rates and under the control of traffic-actuated signals are involved.

A constant yellow interval of 4.5 sec would be able to accommodate the 90th to the 100th percentile requirements at nearly all the intersections. However, because the 95th percentile yellow interval requirement can vary from 3 to 5 sec, the use of a constant yellow interval may not appeal to some traffic engineering agencies. On the other hand, introducing additional variables into a timing procedure in order to account for such a variation would certainly make the resulting procedure impractical. Therefore, it is recommended that simple guidelines be established to assist in the choice of the yellow interval. Such guidelines may evolve on the basis of Equation 15.

At intersections where grades are within ±4 percent, the change interval requirements are strongly and linearly correlated with the vehicle clearance time. Therefore, simple regression equations such as Equations 6 and 7 can adequately serve as a basis for timing design. The ITE's proposed recommended practice lacks a sound theoretical basis and is unnecessarily tedious to apply.

Equation 17 provides a convenient and logical tool for determining the all-red interval requirements. The coefficients of this and other regression equations presented previously can be modified if additional data become available. The existing data base can be enhanced in several respects. Of particular interest are vehicular movements with mean approach speeds exceeding 32 mph, intersections with clearance widths of more than 130 ft, and intersections where grades are steep.

ACKNOWLEDGMENT

Parts of the analyses and the data collection efforts described in this paper are supported by the National Science Foundation in connection with the development of a simulation model.

REFERENCES

1. *Transportation and Traffic Engineering Handbook.* Institute of Transportation Engineers, Washington, D.C., 1982.
2. F. B. Lin. Timing Design of Signal Change Interval. In *Transportation Research Record 1069*, TRB, National Research Council, Washington, D.C., 1986, pp. 46–50.
3. 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.
4. P. S. Parsonson and A. J. Santiago. Design Standards for Timing the Traffic-Signal Clearance Period Must be Improved to Avoid Liability. In *Compendium of Technical Papers*, Institute of Transportation Engineers, 1980, pp. 67–71.
5. ITE Technical Committee 4A-16. Determining Vehicle Change Intervals. Proposed Recommended Practice. *ITE Journal*, May 1985, pp. 61–64.
6. M. S. Chang, C. J. Messer, and A. J. Santiago. Timing Traffic Signal Change Intervals Based on Driver Behavior. Presented at the 64th Annual Meeting of the Transportation Research Board, Washington, D.C., Jan. 1985.
7. R. H. Wortman and T. C. Fox. Reassessment of the Traffic Signal Change Interval. In *Transportation Research Record 1069*, TRB, National Research Council, Washington, D.C., 1986, pp. 62–68.
8. *Uniform Vehicle Code and Model Traffic Ordinances.* The Michie Company, National Committee on Uniform Traffic Laws and Ordinances, Revised 1968, Charlottesville, Va., 1968.

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

A Behavioral Approach to Risk Estimation of Rear-End Collisions at Signalized Intersections

DAVID MAHALEL AND JOSEPH N. PRASHKER

A conceptual approach to estimating the risk of rear-end collisions at a signalized intersection is presented. It is argued that the creation of a large option zone increases the range of the indecision zone, the direction implication of which is an increase in the risk of rear-end collisions. With the aid of field data collected for two warning intervals (3 and 6 sec) before the red light, a large option zone is shown to increase the variance underlying the stopping probability curve, and thus to determine a larger range for the indecision zone. Data from urban intersections support the basic argument that a long warning period causes a significant increase in the number of rear-end collisions.

At present, a general consensus appears to exist in the literature about the effect of traffic signals on rear-end collisions. Most of the researchers apparently have concluded that signalizing an intersection significantly increases the number of rear-end collisions. For example, in a sample of 34 urban intersections Hakkert and Mahalel *(1)* found that after the introduction of a traffic signal control, the annual number of rear-end collisions increased from 33 to 77. In a similar study of 31 intersections in Milwaukee, Short et al. *(2)* found an increase of 37 percent in the number of such accidents. King and Goldblatt *(3)* observed the same phenomenon of increased rear-end collisions in a statistical analysis of U.S. accident data nationwide.

In addition to the fact that the number of rear-end accidents increases after the introduction of traffic signals at an intersection, it is typical that the highest number of accidents at signalized intersections are rear-end collisions. A statistical analysis conducted by the author of almost all signalized intersections in Israel (600) indicated that, over a 2-year period (1983 to 1985), about 39 percent of all accidents were rear-end collisions, compared with about 27 percent that were right-angle collisions. A similar result was identified by Hanna et al. *(4)* in a study of signalized-intersection accidents in rural communities in Virginia. They found that 43 percent of all accidents were rear-end collisions and 37 percent were right-angle collisions.

In spite of the fact that rear-end collisions at signalized intersections are significantly more frequent than right-angle collisions, the former have not received much attention either in the literature or in practice. The reason for the interest in right-angle collisions is their relatively high severity. The general belief is that an improvement in existing procedures defining the change interval will gradually reduce the number of right-angle collisions. The term "change interval" means the sequence of intervals at a signalized intersection that occurs from the moment the green light ends for one direction and a green light begins for the conflicting approach. The change interval might consist of various combinations and proportions of yellow and red for the conflicting approaches. Despite the consensus referred to at the outset, much of the research to date that is concerned with accidents at signalized intersections concentrates on proper design procedures for the change interval to reduce the number and severity of right-angle collisions.

Because of the large number of rear-end collisions, and in spite of their relatively low severity, the authors believe this type of accident deserves more attention. The purpose of this paper is to present a conceptual approach to estimating the risk of rear-end collisions at a signalized intersection. This work is based on a behavioral analysis of drivers approaching an intersection when the yellow light appears at the end of the continuous green light. An analysis will be made of the influence of the length of the warning interval—the interval between the continuous green and the continuous red—on the probability of a rear-end collision. The theoretical analysis is then supported by field observations carried out under controlled conditions at several signalized intersections.

RISK-GENERATING PROCESS OF A REAR-END COLLISION

Most rear-end collisions at a signalized intersection occur when two successive drivers approaching the intersection make conflicting decisions when the yellow light appears. A high risk of a rear-end collision will exist if the first driver decides to stop while the second one wants to cross the intersection. When the collision actually occurs, it is reasonable to assume that the second driver did not anticipate the stopping decision of the driver in front, and thus could not react in time to prevent the accident. The highest probability of a rear-end collision exists when the probability of two successive drivers reaching conflicting decisions about whether to cross or stop is the highest.

The probability for conflicting decisions is a function of the distance or time of the two drivers from the intersection when the yellow light appears. This probability of conflicting decisions can be derived from a stopping probability function that describes the probability of stopping when the yellow light

Civil Engineering Faculty and members of the Transportation Research Institute, Technion-Israel Institute of Technology, Haifa, Israel.

appears as a function of the distance from the intersection.

Let $P(x)$ be the probability of stopping when a driver is at distance x from the intersection when the green phase ends. The probability of deciding to cross the intersection will then be $1 - P(x)$. Note that this function (Figure 1a) represents a realization of Bernoulli trials carried out at various distances from the stop line the moment the green light ends. The probability of these trials changes as a function of distance; the probability of stopping is high when the driver is relatively far from the intersection, and low when the driver is close.

The probability of two drivers' reaching conflicting decisions about whether to pass or stop will be highest when the expression $P(x) * [1 - P(x)]$ obtains the maximum value. This happens when $P(x) = 0.5$. Figure 1b shows this probability, the value of which becomes lower as the distance to the stop line increases or decreases.

The zone around the point at which the stopping probability has a value of 0.5 is the zone in which it is most difficult for the driver to reach a decision on the proper action when the green light ends.

In practice *(5–7)*, it is customary to describe the area between the 10th and 90th percentiles of the stopping probability function as an indecision zone. An example of the implementation of the concept of the indecision zone is found in a work by Parsonson *(6)*. He suggested placing in this zone a detector loop whose purpose would be to prevent, in unsaturated cycles, a situation in which a driver is caught in the indecision zone at the beginning of the yellow light.

A necessary condition for the occurrence of a rear-end collision is the presence of vehicles in the intersection approach when the yellow light appears. The probability of a rear-end collision increases when the number of vehicles in the indecision zone increases. The actual number of rear-end accidents is thus a function of two factors:

1. Traffic volume—the larger the volume of vehicles in the approach to the intersection, the higher the probability that vehicles will be located in the indecision zone when the green light ends.
2. The range of the indecision zone—the larger the indecision zone, the higher the probability that vehicles will be located in the zone when the green light ends; the range of the indecision zone depends on the value of the variance of the random variable that generated the stopping probability function.

The characteristics of this random variable and the stopping probability function is discussed in the section entitled The Relation Between the Option and Indecision Zones. At this point, it will be sufficient to show the effect of this variance. Figure 2 shows two stopping probability functions that differ in their underlying variance. It is easy to see that the larger this variance, the larger the indecision zone.

The behavioral aspects of the causes of rear-end collisions have been discussed, and now deterministic normative methods to analyze the intersection-approach problem will be discussed next.

DILEMMA AND OPTION ZONES

In many studies *(8–11)* concerned with the events occurring in the approach to a signalized intersection, the phenomena are analyzed through the use of dilemma and option zones. These zones are defined by equations that are based on the normative behavior of a reasonable driver when the yellow light flashes.

Drivers who are located in the dilemma zone at the end of

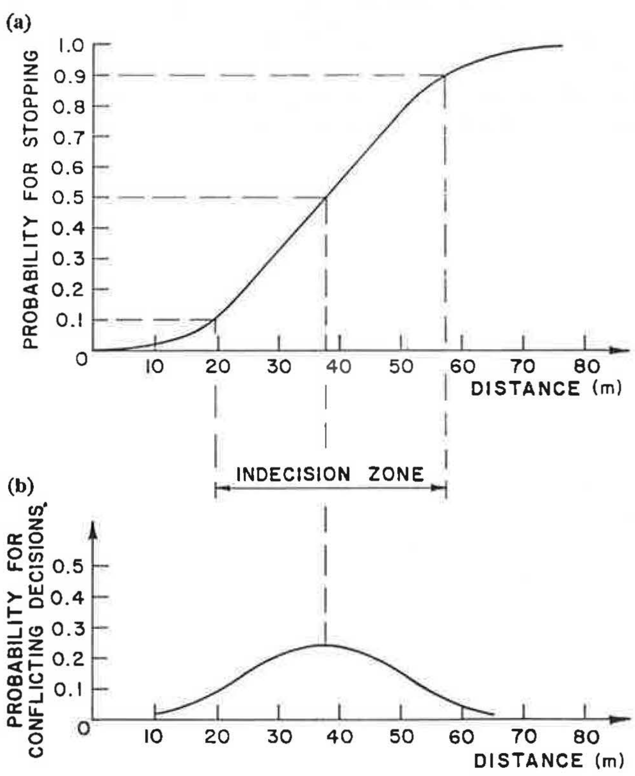


FIGURE 1 Hypothetical stopping probability function and the probability for conflicting decisions.

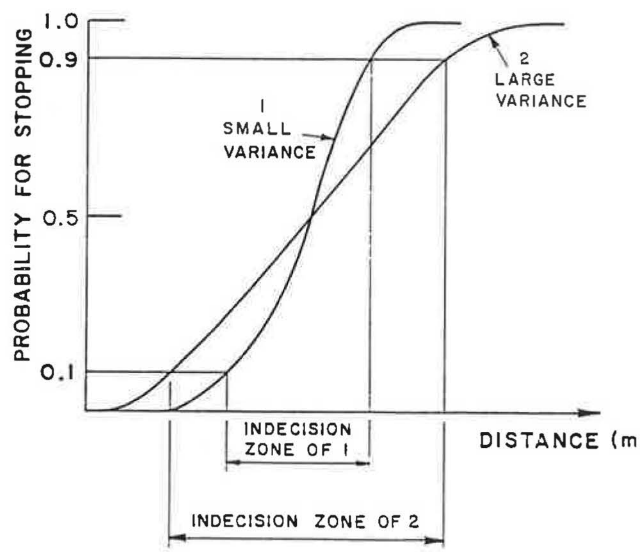


FIGURE 2 Two hypothetical stopping probability functions in which $P(x) = 0.5$ occurs at the same x and that have different underlying variances.

the green light can neither stop their vehicles before the stop line nor cross the line before the light turns red. Drivers who are in the option zone when the light turns yellow can either stop their vehicles at the stop line or cross it before the light turns red. The ability of a driver to cross the stop line or to stop before it is based on deterministic normative values. It is usually assumed that deceleration takes place at a rate of about 10 ft/sec^2 and that when an attempt is made to cross the intersection, the driver will continue at a constant speed or accelerate at a rate of 5 ft/sec^2 *(9)*. Figure 3 shows the shape of the dilemma and option zones as a function of approach speed.

The importance of the definition of dilemma and option zones lies in a normative ability to analyze and judge various actions taken by drivers at an intersection approach. For example, May defined a risk-measurement factor based on the events occurring in dilemma and option zones *(9)*. It is important to realize, however, that these zones describe, under normative deterministic assumptions, what a driver can do in each zone. They do not describe what a driver will actually do, not even in the stochastic sense. Thus, it can be concluded that the dilemma and option zones are tools of diagnostics or analysis; they are not, and cannot describe, the actual behavior of drivers.

In many of the studies carried out following the work by Gazis et al. *(8)*, special emphasis was placed on reducing the size of the dilemma zones. The motivation behind this objective was to lessen the risk of right-angle collisions. The manifestation of this school of thought is the ***Proposed Recommended Practice for Determining Vehicle Change Interval*** by ITE *(12)*. In these guidelines, the proposed speed approach for determining the length of the yellow light is the 85th percentile of the actual speed distribution or of the posted speed limit. This recommendation indicates that the tendency is to use a relatively high approach speed to reduce the size of the dilemma zone. By doing so, there is a smaller chance that a driver who is unable to stop before the stop line when the red signal lights up will eventually cross during the red light. This, of course, is in line with legal attitudes as expressed in traffic laws.

The direct implication of determining the length of the yellow light according to the relatively fast drivers is to create a large option zone for the slower drivers. This option zone provides slow drivers a relaxed decision situation because whether they decide to stop or cross the intersection, they can do so within the legal time frame. Although this situation may be desired by the individual driver, it has serious implications at the system level.

The option zone, by definition, is an area in which either decision—to stop or to cross—is legitimate; thus, a high proportion of conflicting decisions may be expected by the various drivers located in this zone. The high proportion of possible conflicting decisions by itself creates a high potential for rear-end collisions. To demonstrate this contention, imagine that a stop sign is considered by some drivers to be a recommendation to stop and by others a recommendation to cross the intersection. This situation, by its very nature, will create conflicts and, thus, rear-end collisions. The hypothetical situation is analogous to the interpretation of the option zone advanced here.

In the next sections is an analysis, through field data, of whether increasing the option zone influences the indecision zone as defined in this paper.

RELATION BETWEEN THE OPTION AND INDECISION ZONES

The implications of increasing the size of the option zone on the size of the indecision zone were analyzed under controlled conditions at four urban signalized intersections in Tel Aviv. The events at those intersections were twice recorded on film with a cine camera—once with a small option zone and once with a larger zone. At each intersection, the results of the experiment were first recorded in the situation in which signals always operated (either with or without flashing green). The mode of operation at these intersections was changed to the other mode. Appropriate announcements were made on the

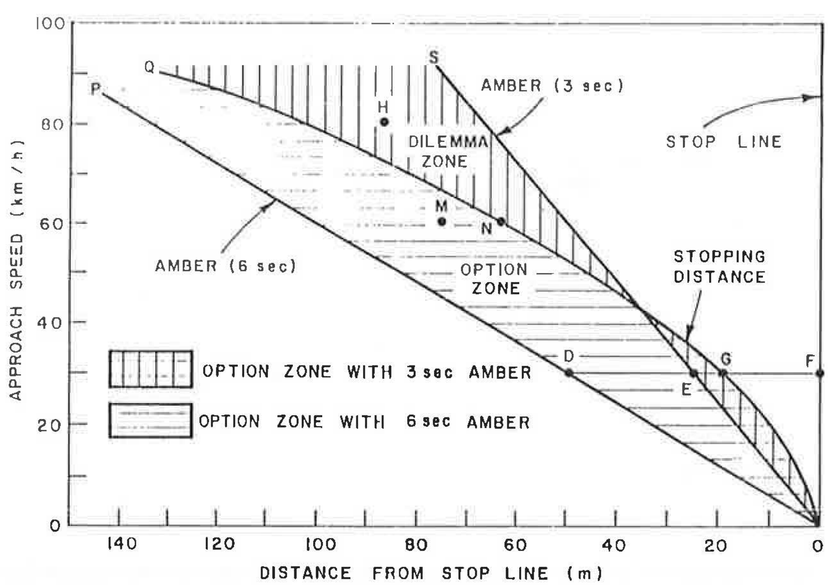


FIGURE 3 Options and dilemma zones for two warning periods.

radio, and the second set of experiments in the new mode of operation was made 1 to 2 months later. Thus, it can be assumed that drivers adjusted to the new operation mode of signals. A detailed description of the data is given by Becker *(13)*.

The increase in the option zone was achieved, not by lengthening the yellow light, but by substituting the last 3 sec of the green light with a flashing green. The flashing green was not new to the Israeli driver because it had been used in Israel for years at most interurban signalized intersections and at those urban intersections that had high approach speeds. Thus, it can be assumed, as was done in this study, that the flashing green is perceived by the Israeli driver mostly as an extension of the warning period, to indicate the close appearance of the red light. As shown in Figure 3, the increase in the warning period significantly increases the option zone.

Data Description

The basic characteristics of the four intersections included in the sample are given in Table 1. The size of the sample is the number of vehicles at the end of the green light that were actually exposed to a stopping or crossing decision. The sample does not include vehicles forced to stop by vehicles in front.

The first three intersections may be seen to be characterized

TABLE 1 BASIC DATA OF THE FOUR INTERSECTIONS INCLUDED IN THE STUDY

Intersection No.	Characteristic	Warning Period [1]	
		3 sec.	6 sec.
1	No. of cycles	75	67
	Sample size (veh)[2]	52	47
	Volume (VPH)	320	330
	Average speed (km/h)[3]	31.7	37.8
	Speed variance	76	147
2	No. of cycles	80	83
	Sample size (veh)	256	341
	Volume (VPH)	950	865
	Average speed (km/h)	22.1	21.4
	Speed variance	95	72
3	No. of cycles	73	68
	Sample size (veh)	255	239
	Volume (VPH)	1317	1404
	Average speed (km/h)	37.1	34.7
	Speed variance	98	111
4	No. of cycles	42	47
	Sample size (veh)	60	131
	Volume (VPH)	Missing Data	1065
	Average speed (km/h)	59.4	63.0
	Speed variance	156	85

[1] The 6 sec. warning period is composed of 3 sec. flashing green and and 3 sec. yellow.

[2] The sample size refers to the number of vehicles at the end of the green. The vehicles preceded by stopped vehicles are not included.

[3] Speed of vehicles at the end of the continuous green.

by a low approach speed (20 to 40 km/h), and the fourth by a higher approach speed (60 km/h). The data were collected in two situations:

1. *Three-second warning*: a yellow light appeared for 3 sec after the continuous green light and was followed by the red light.
2. *Six-second warning*: a flashing green light of 3-sec duration appeared after the continuous green, followed by a yellow light for 3 sec, and then the red light. All together, the warning period lasted for 6 sec.

For each vehicle, the data included its position at the end of the continuous green light, its speed, and its deceleration rate if it stopped.

The Model

As mentioned previously, the indecision zone is derived from the stopping probability function, which was estimated according to Sheffi and Mahmassani *(7)*. According to that model, a variable T is defined as the driver's perceived time to the stop line at the beginning of the warning period (the end of the continuous green). Because of differences in driver perception, T was assumed to be a random variable with normal distribution; that is,

$$T \sim N(t, \sigma_T^2)$$

In addition, it was assumed that if T is less than a critical value, T_{cr}, a driver would decide to cross the intersection; otherwise, he would decide to stop. T_{cr} was also assumed to be a random variable:

$$T_{cr} \sim N(t_{cr}, \sigma_{cr}^2)$$

Consequently, the probability for stopping at the end of the green light is

$$P_{\mathrm{STOP}}(T) = P(T_{cr} \leq T) = \Phi\left(\frac{t - t_{cr}}{\sigma}\right)$$

where $\sigma = \sqrt{\sigma_T^2 + \sigma_{cr}^2 - 2\sigma_{T,cr}}$ and $\Phi(\bullet)$ denotes the standard cumulative normal function.

The two parameters of the model (t_{cr}, σ) were estimated with a program called CHOMP (Choice Modeling Program), which is used to estimate multinomial probit models.

Results

The expected value of the critical distance (t_{cr}) defines the distance from the intersection in which 50 percent of the drivers at the beginning of the warning period will eventually stop and 50 percent will cross. The majority of drivers located between the stop line and the expected value of the critical distance at the beginning of the warning period (yellow or flashing green) will decide to cross; at longer distances than the critical, most will decide to stop. It is reasonable to assume that the expected value of the critical distance will be between the stopping distance and the distance a vehicle can travel during the warning period.

From the results given in Table 2, it can be seen that in all the samples, the influence of lengthening the warning period is expressed in longer expected critical distances (t_{cr}). The warning period in the present study was changed from 3 to 6 sec, so a 3 sec change in t_{cr} demonstrates that, in terms of expected time from the intersection, the stopping decisions of drivers undergo no alteration. In other words, even in cases of a longer warning period, it is the beginning of the red light that determines the behavior of drivers rather than the length of the warning period.

The actual increase in the length of the warning period was 3 sec. However, from Table 2 it appears that the increase in t_{cr} was not always exactly 3 sec. When the increase in t_{cr} is less than 3 sec there will be a higher number of stopping decisions as compared with an increase of exactly 3 sec. This means that there exists a group of drivers who, when at the same location, decide to cross with short warning intervals, but stop with long warning intervals when the solid green ends. In cases in which the change in t_{cr} is greater than 3 sec, there exists a group of drivers who could stop with a short warning interval, but a long warning period causes this group to cross.

At Intersections 1 and 2 (see Table 2), the change in the expected critical distance is very close to 3 sec (2.85 and 3.19 sec, respectively). At Intersection 3, the change is 3.96 sec, and at Intersection 4, 2.18 sec. Considering the limitations of the sample size and the small number of intersections, it may be concluded that, in terms of expected critical distance, a significant change in driver behavior does not occur from lengthening the warning period.

The variance underlying the stopping probability function increased significantly at the first three intersections; at Intersection 4, the variance changed only slightly (Table 2). For example, the variance at Intersection 1 increased from 0.31 to 6.14. The direct implication of the increased variances is an increase in the range of the indecision zone. As can be seen, this increase is about 90 to 350 percent. For example, at Intersection 1, the range of the indecision zone increased from 1.42 to 6.34 sec. The stopping probability curves of Olson and Rothery *(14)*, which were estimated during two yellow durations, also show a tendency for long-range indecision zones during a long yellow light.

Despite the fact that the range of the indecision zone did not change at one intersection only, it appears that this result is not random. As previously mentioned, Intersection 4 is characterized by a relatively high approach speed. Stopping distances, therefore, are longer here than at the other intersections. This means that drivers who can stop are situated farther from the intersection. It is reasonable to assume that a driver's temptation to cross when at a short distance from the stop line is higher than when at a long distance; thus, a driver is more likely to stop at option zones with high-speed approaches than at option zones with low approach speeds. This finding is related to the hypothesis advanced by Mahalel and Zaidel *(15)* that drivers' stopping decisions are more strongly influenced by their distance from the stop line than by their approach speeds. At low approach speeds, drivers are likely not to stop even if

TABLE 2 INDECISION ZONE BOUNDARIES FOR TWO WARNING INTERVALS

Intersection	Warning Interval[1] (sec)	$\hat{t}_{cr}$ (sec)	$\hat{\sigma}^2$	Indecision Zone Boundaries[2]		
				Inner boundary (sec)	Outer boundary (sec)	Total length (sec)
1	3	2.42	0.31	1.71	3.13	1.42
	6	5.27	6.14	2.10	8.44	6.34
2	3	1.45	0.87	0.26	2.64	2.38
	6	4.64	3.16	2.36	6.91	4.55
3	3	3.12	0.53	2.19	4.05	1.86
	6	7.08	4.98	4.22	9.94	5.72
4	3	4.46	1.39	2.95	5.97	3.02
	6	6.64	1.42	5.11	8.16	3.05

[1] A warning interval of 3 sec. consists of 3 sec. amber. A warning interval of 6 sec. consists of 3 sec. flashing green and 3 sec. amber.

[2] The inner and outer boundaries are the values of the 10th and 90th percentiles respectively of the stopping probability function.

they can; at high approach speeds, whenever drivers can stop they do so with high probability. Evidence that drivers' decisions reflect a higher sensitivity to distance than to speed may also be found in Chang et al. *(16)*.

DISCUSSION OF FINDINGS

The relationship between the option zone and the range of the indecision zone has been discussed in this paper. Empirical evidence demonstrated that with a low approach speed, the resulting increased option zone causes a significant increase in the indecision zone.

The calculations of conflicting decisions were made with the assumption of independency behavior between consecutive drivers. This simplified assumption ignores the possibility of dependency as a result of car-following behavior. However, it is logical to assume that the monotone relationships between the stopping probability and rear-end accidents will exist, even if dependency will be taken into consideration in a more sophisticated model.

In an analysis of the risk of rear-end collisions (see section entitled Risk-Generating Process of a Rear-End Collision), it was assumed that this risk might increase as a result of an increase in the range of the indecision zone. This hypothesis thus indicates the possibility that the addition of a flashing green light as an extra warning period might increase the number of rear-end collisions. Various studies *(17–19)* have revealed that the number of rear-end collisions at urban intersections with a flashing green signal is significantly higher than at other signalized intersections. This important finding corroborates the hypothesis regarding the relationship between the range of the indecision zone and the risk of rear-end collisions.

The operational implication of this relationship is that additional research should be conducted to find the pattern of warning intervals that will minimize the range of the indecision zone. In other words, ways should be found to shape the stopping probability function to be as close as possible to a step function. In that way, the number of rear-end collisions might decrease.

REFERENCES

1. A. S. Hakkert and D. Mahalel. The Effect of Traffic Signals on Road Accidents with Special Reference to the Introduction of a Blinking Green Phase. *Traffic Engineering Control*, Vol. 19, No. 5, 1978.
2. M. S. Short, G. A. Woelel, and C. J. Chang. Effects of Traffic Signal Installation on Accidents. *Accident Analysis and Prevention*, Vol. 14, No. 2, 1982, pp. 135–145.
3. G. F. King and R. B. Goldblatt. Relationship of Accident Patterns to Type of Intersection Control. In *Transportation Research Record 540*, TRB, National Research Council, Washington, D.C., 1975, pp. 1–12.
4. J. T. Hanna, T. E. Flynn, and W. K. Tyler. Characteristics of Intersection Accidents in Rural Municipalities. In *Transportation Research Record 601*, TRB, National Research Council, Washington, D.C., 1976, pp. 79–82.

5. C. V. Zegeer. *Effectiveness of Green Extension System at High Speed Intersections.* Research Report 472. Bureau of Highways, Kentucky, Department of Transportation, Louisville, 1977.
6. P. S. Parsonson. Signalization of High Speed Isolated Intersection. In *Transportation Research Record 681*, TRB, National Research Council, Washington, D.C., 1978.
7. Y. Sheffi and H. Mahmassani. A Model of Driver Behavior at High Speed Signalized Intersections. *Transportation Science*, Vol. 15, No. 1, Feb. 1981.
8. D. Gazis, R. Herman, and A. Maradudin. The Problem of the Amber Signal Light in Traffic Flow. *Traffic Engineering*, July 1960.
9. A. D. May. Clearance Interval at Traffic Signals. In *Highway Research Record 221*, HRB, National Research Council, Washington, D.C., 1968, pp. 41–71.
10. H. H. Bissell and D. L. Warren. The Yellow Signal is not a Clearance Interval. *ITE Journal*, Feb. 1981.
11. *Transportation and Traffic Engineering Handbook.* Institute of Traffic Engineers, Prentice Hall Inc., Englewood Cliffs, N.J., 1976.
12. Proposed Recommended Practice Determining Vehicle Change Intervals. ITE Technical Committee 4A-16. *ITE Journal*, May 1985.
13. M. Becker. *A Study of the Effect of a Flashing Green Phase in Traffic Signals at Urban Intersections.* Publication 71/7. Road Safety Centre, Technion Research and Development Foundation Ltd., Haifa, Israel, 1971 (in Hebrew).
14. P. L. Olson and R. W. Rothery. Driver Response to the Amber Phase of Traffic Signals. *Operations Research*, Vol. 9, No. 5, 1961.
15. D. Mahalel and D. M. Zaidel. A Probabilistic Approach for Determining the Change Interval. Presented at 65th Annual Meeting of the Transportation Research Board, Washington, D.C., 1986.
16. M. S. Chang, C. J. Messer, and A. J. Santiago. Timing Traffic Signal Change Intervals Based on Driver Behavior. Presented at the 64th Annual Meeting of the Transportation Research Board, Washington, D.C., 1985.
17. D. Mahalel and D. M. Zaidel. Safety Evaluation of a Flashing Green Light in Traffic Signal. *Traffic Engineering and Control*, Vol. 26, No. 2, 1985.
18. A. S. Hakkert and D. Mahalel. The Effect of Traffic Signals on Road Accidents with Special Reference to the Introduction of a Blinking Green Phase. *Traffic Engineering and Control*, Vol. 19, No. 5, 1978.
19. I. Hocherman and J. Prashker. *Identification of High Risk Intersections in Urban Areas.* PTRC Summer Annual Meeting, Brighton, U.K., 1983.

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

EVIPAS: A Computer Model for the Optimal Design of a Vehicle-Actuated Traffic Signal

A. G. R. BULLEN, NORMAN HUMMON, TOM BRYER, AND ROSLI NEKMAT

The EVIPAS model described is a computer program designed to analyze and optimize a wide range of intersection, phasing, and controller characteristics of an isolated fully actuated traffic signal. It will evaluate almost any phasing combination available in a 2- to 8-phase NEMA type controller and similar phasing structures for a Type 170 controller. The model will provide optimum timing settings for pretimed, semi-actuated, fully actuated, or volume-density control using a variety of measures of effectiveness chosen by the user. A wide range of geometric features, phasing alternatives, and detector layouts can be evaluated. EVIPAS combines a user friendly input module with a multivariate gradient search optimization module and an event-based intersection microsimulation. It has been field-tested and validated and replicates well-observed vehicle and signal behavior. The model is programmed in Fortran 77 and currently can run on VAX 8600 and IBM 3080 mainframes.

In recent years traffic signal design has been facilitated by the increasing availability of computer software for signal timing analysis. Most of the models available, however, are calibrated for pretimed signals. The *1985 Highway Capacity Manual (1)* provides a set of capacity analysis procedures for signalized intersections that are heavily dependent on a signal timing and phasing plan. Further, a large number of the new signals being installed in North America are fully vehicle actuated with many of these using volume-density control. Therefore, a clear need exists for software that will provide optimal design for vehicle-actuated traffic signals.

The software that is currently available has only limited applicability. The only single intersection model *(2)* that optimizes design parameters is SOAP84. This model depends heavily on the approach of Webster *(3)*, which is mainly for pretimed signals. Although SOAP84 does provide some assistance for dealing with vehicle actuation, it does not attempt to provide a complete analysis capability for the many options that are available.

The TEXAS model *(4)* is not widely circulated. It is a microsimulation of an intersection with vehicle-actuated signals, but provides no direct optimization capability. The model is rather slow, and it is not clear how well it deals with all of the individual timing parameters for fully actuated volume density control.

A. G. R. Bullen, R. Nekmat, Department of Civil Engineering, and N. Hummon, Department of Sociology, University of Pittsburgh, Pittsburgh, Pa. 15261. T. Bryer, Pennsylvania Department of Transportation, 215 Transportation and Safety Bldg., Harrisburg, Pa. 17120.

NETSIM is a network traffic microsimulation that has a detailed vehicle-actuated signal capability for individual intersections *(5)*. This model has no direct optimization capability and is also slow. It is primarily intended for the analysis of area control type problems.

The EVIPAS model described in this paper is able to analyze and optimize a wide range of intersection geometric configurations, phasing, and controller characteristics of a fully vehicle-actuated isolated traffic signal. It will evaluate almost any phasing combination available in a one- or two-ring NEMA type controller and similar phasing structures for a Type 170 controller. It will provide the optimum timing settings for pretimed, semi-actuated, fully actuated, or volume-density control by using a variety of measures of effectiveness chosen by the user. These include delay, fuel consumption, other operating costs, and emissions. A wide range of geometric features, phasing alternatives, and detector layouts can also be evaluated.

The EVIPAS model has a user friendly input module and is currently programmed for the VAX 8600 and IBM 3090 mainframes.

BACKGROUND

VIPAS was a model originally developed at the Pennsylvania Department of Transportation (PennDOT) in Harrisburg, Pennsylvania by Tom Bryer and programmed by John Breon *(6)*. The department realized the need for a model to optimize actuated signal design and also provide an estimation of the economic benefits of installing actuated traffic signals.

The major component of VIPAS was a microsimulation of a signalized intersection. The simulation was a second-by-second vehicle scanning procedure using the car-following algorithms of the Federal Highway Administration INTRAS freeway simulation *(7)*. For vehicle queues and queue discharge from the stop line, more efficient flow discharge models substituted for the individual vehicle scanning process.

An unusual characteristic of the model was the randomly generated vehicle arrivals for multilane approaches. It has been well established that the total arrival pattern of a multilane approach is not just the simple sum of the random distributions on the individual approaches due to the correlation between vehicles across lanes. To overcome this problem VIPAS used specific multilane arrival distributions calibrated from test runs by FHWA on the INTRAS simulation. These distributions were stored in VIPAS as the inverse distribution functions in the form of nth degree polynomials.

Other features of the initial VIPAS model included four vehicle types, including car, bus, truck, and semi-trailer, with acceleration rates particular to each vehicle type. The acceleration for each vehicle type had two ranges with a higher acceleration below a given threshold speed. Similarly, for deceleration, all vehicles coast at a low deceleration between their desired speed and a deceleration threshold of 0.9 times the desired speed. Below this threshold a greater deceleration is imposed. The desired speeds for the vehicles are generated randomly by the normal distribution.

Various measures of effectiveness were available to the user. These include total vehicle delay, stopped delay, person delay, fuel consumption, total operating costs, and vehicle emissions.

The VIPAS model was implemented on the department's IBM 3081 computer. Operational use of the model revealed a number of difficulties including the following:

1. It dealt only with a restricted set of geometric and phasing configurations.
2. It allowed only one detector per approach lane.
3. The model required further field verification.
4. There was a desire to make it more user friendly.
5. The total was very large and rather slow, which inhibited its use by field engineering staff in the department.

To enhance the capabilities of the VIPAS model and correct any deficiencies, the University of Pittsburgh was awarded a research project by the PennDOT Office of Research in 1985. The objectives of this project were to expand and generalize the capabilities of VIPAS, carry out field studies for calibration and validation, provide a user friendly input structure, and make the overall model smaller and faster.

THE NEW MODEL

An analysis of the structure of the VIPAS model indicated that the simulation could not be easily generalized to cover the required broad range of traffic and signal conditions, and the optimization methodology was not suitable for full multivariate situations. Consequently, a new optimization algorithm and a new intersection simulation were designed and programmed. The original VIPAS traffic characteristics and vehicle generation routines were combined with these new models to give the enhanced version EVIPAS.

The new EVIPAS model consists of five major modules:

1. An input module that provides a user friendly interface for the user: INPROC,
2. A generation module that generates all vehicle and pedestrian arrivals and their characteristics: GENRAT,
3. An optimization module that finds the optimum settings of the selected timing variables: OPTSIM,
4. An intersection simulation that provides the function calls for the optimization: PROCES, and
5. An output module: OUTPUT.

Input Processing: INPROC

The purposes of the input processing routines are to

1. Provide a user friendly environment whereby appropriate data files can be created, updated, and edited using an interactive or batch mode; and
2. Process or transform user data into error free, compiled data for use in the VIPAS simulation and optimization processes.

INPROC, which is designed to be used independently, helps users to create new data files, update existing files, and edit appropriate data elements by using an interactive or batch mode. Each data file is used for a specific project. INPROC then creates internally two compiled files that will be read by GENRAT and PROCES. This ensures that the optimization and simulation runs are on error-free data.

VIPAS requires three data sets per data file: (a) intersection characteristics, (b) signalization characteristics, and (c) traffic characteristics.

INPROC will guide users through all three data sets from data element to data element in logical sequence or at the users' option to edit, review data, or seek help. Given a strict and inflexible FORTRAN 77 programming environment, this liberal input philosophy is made possible by a format-free, input interface routine (FREFRM) whereby all users' input is received and assessed for its validity in terms of the data elements requested. Valid inputs are those within predetermined upper and lower bounds. This strategy is used to filter out outliers and unqualified inputs. A default value is assigned with the user's approval when an invalid input is encountered. FREFRM provides the primary mechanism to filter inputs and to achieve some degree of user friendliness while still operating in the FORTRAN 77 programming environment. FREFRM adopts suggestions by Wright *(8)* in terms of man-machine interfacing in the FORTRAN environment.

For intersection characteristics, users are requested to provide information pertaining to the physical features of the study intersection, including (a) number of approaches (maximum of 5); (b) number of lanes in each approach (maximum of 5); (c) detector locations (combinations of presence or passage detectors, or both, with up to three per lane); and (d) saturation flows adjusted only for width and gradient factors according to the *1985 Highway Capacity Manual (1)*. For a particular case study, the intersection characteristics may be kept constant or varied to test alternative physical configurations. EVIPAS, however, simulates and optimizes for a given set of intersection characteristics.

EVIPAS is designed to simulate and optimize fully actuated, semi-actuated, volume-density, and pretimed signalization with or without pedestrian actuation. For simulation purposes, users are required to provide various timing parameter values for each phase defined. For an actuated signal this would include the initial intervals, the unit extensions, the maximum intervals, and the yellow and all-red clearance intervals. For optimization purposes, users are not required to provide the timing for those parameters they are choosing to optimize. VIPAS will create its own timing parameters as a starting point before determining the optimum values. Users may provide upper or lower bounds, or both, for the variables being optimized.

The bulk of the data input requirement is in the definition of the traffic characteristics. INPROC can handle a week's data that has been segmented or separated into periods of similar traffic characteristics, such as morning peak-hours during weekday, and weekend traffic. For each period, users are to provide parameters that will typically describe the traffic characteristics during that period including (a) volume per lane, (b) traffic composition, (c) traffic turning movements, (d) average speeds, and (e) pedestrian counts. Periods may be linked together and can be assigned different weighting factors.

The final stage of INPROC is to check and compile user inputs into error-free data. A special routine, COMPILE, checks for errors and inconsistencies in user data hierarchically from intersection level to approach level and from approach level to lane level. In addition, data elements are checked for errors and inconsistencies at the same level. If any data element is found in error, the user is prompted by COMPILE to make the appropriate correction.

Vehicular and Pedestrian Generation: GENRAT

The purpose of GENRAT is to generate stochastically the traffic and pedestrian arrivals as defined by the traffic characteristics in INPROC for each period of study. Vehicles are generated randomly at the source of each approach, which is a predefined point upstream of the intersection stop line where traffic flow is not influenced by the intersection. Pedestrian arrivals are randomly generated at the push-button. Vehicle and driver characteristics, which include (a) arrival time, (b) vehicle type, (c) driver type, (d) source lane, (e) turning direction, (f) speed, and (g) follower, are assigned randomly to each vehicle generated.

For headways for low traffic volumes of less than 40,100 and 200 vph for one-, two-, and three-lane approaches, a negative exponential distribution is assumed. For one-lane approaches with higher volumes, Kell's composite exponential distribution is used *(9)*. For multilane approaches with higher volumes, the probability distribution function calibrated from INTRAS is used. Special care is taken to ensure logical sequence and proportionality of arrivals per lane at the source, especially in multilane approaches.

Other vehicle characteristics such as vehicle type, driver type, source lane, and turning direction are generated by discrete uniform distributions, whereas vehicle speeds use the normal distribution and pedestrian arrivals use the negative exponential distribution.

The output of GENRAT is a set of vehicles stored in a data file, with assigned characteristics for each approach during the period of study. Users have the option of checking the statistics of the GENRAT output by using the GENSTAT routine, which is a support module that computes statistics for the arrival data generated by the GENRAT module. For each lane and period, GENSTAT computes the mean and standard deviation of headways; median headway; minimum and maximum headway; the order statistics for the quantiles 1, 5, 10, 25, 75, 90, 95, and 99 percent; mean and standard deviation of the vehicle speed; vehicle-type frequency distribution; turn-direction frequency distribution; and driver-type frequency distribution.

Optimization: OPTSIM

As with most current traffic models, such as TRANSYT *(10)*, the optimization in the original VIPAS was a sequential univariate procedure. The problem with using these types of methods is the time needed to converge and the fact that the method does not guarantee a local optimum.

The optimization module in the EVIPAS is a multivariate procedure that uses quasi-Newton methods to find the optimal values of the parameters. Numerical procedures compute first and second derivatives of the function for a given vector of parameter values. The values of these derivatives are used to determine the direction and size of steps from one parameter vector to the next and to determine whether the minimum is reached. The algorithms are designed to avoid certain types of local minima, although there exist conditions for which these methods fail to find the optimum solution. In general, however, these methods are among the best available for the solution of this class of problems.

The function for the optimization is the measure of effectiveness (MOE) output of the intersection simulation. To make a function out of a simulation model, three computational problems must be addressed. First, the model must be structured so that a run of the simulation model is functionally dependent only on the values of the vector of input parameters. All other data necessary for the operation of the model must be fixed for all iterations of an optimization run. Second, at the start of each iteration, the model data and starting conditions must be identical. Thus the model must be reinitialized to the same status at the beginning of each iteration. Third,, the computation of the output value must not change from iteration to iteration. Meeting these requirements is an important feature of the EVIPAS structure.

From the perspective of the optimization module, a function is a function. It matters not whether the function has a simple analytic form or is a large simulation model. The optimization module is only concerned with the relation of the parameter vector and the associated output value. The optimization algorithms and numerical methods used in EVIPAS are on pseudocode programs and subroutines reported by Dennis and Schnabel *(11)*.

In implementing this general optimization algorithm, several important choices must be made. The first is the determination of the size of the linear step to ensure that an improvement in the function will occur. If no improvement is found, the algorithm backtracks to determine a better one by fitting a cubic to the last few values and solving the cubic for its minimum.

The second choice concerns how to compute the Hessian matrix, which can be done by finite-difference approximations or secant update methods. The finite-difference approximations require that the function be evaluated many times, whereas the secant update approach does not require that the function be evaluated; instead it solves a set of linear equations using the old Hessian method and the current values of the gradient and parameter vector. Because function evaluations are computationally expensive in this application, the preferred method is the secant update approach as it minimizes the number of times the simulation model must be run.

Nevertheless, for large problems, many function evaluations are required, perhaps approaching several hundred. Fortunately, most signal problems involve a relatively small number of variables and the optimization converges rapidly.

Intersection Simulation: PROCES

The simulation of an intersection under vehicle-actuated control presents a set of traffic movement alternatives that are so complex as to require a microsimulation. However, most microsimulations require considerable computer time—a major disadvantage for a model such as EVIPAS, which requires many function calls in its optimization.

The simulation that has been developed for the EVIPAS has several structural features that are designed to enhance its efficiency. Primarily it is an event-based individual vehicle simulation, with the events being green extensions or green termination. The simulation constructs vehicle trajectories in space-time according to a linear car-following model, which is related to the model used in the INTRAS freeway simulation. The car-following algorithm has been reformulated, however, to provide a more realistic handling of driver reaction times. Each trajectory consists of a series of nodes that represent changes in acceleration. The linear car-following model is

$$a(n) = b[v(n) - v(n-1)] \tag{1}$$

where

$a(n)$ = acceleration of the nth vehicle at time $t + T$,
$v(n)$ = speed of the nth vehicle at time t,
$v(n-1)$ = speed of the $(n-1)$th vehicle at time t,
T = driver reaction time, and
b = a coefficient.

This formulation leads to a headway condition

$$x(n) = x(n-1) + M + kv(n) \tag{2}$$

where

$x(n)$ = distance coordinate of vehicle n at time t,
$x(n-1)$ = distance coordinate of vehicle $n-1$ at time t,
M = minimum stopped distance headway between the vehicles, and
k = driver reaction time = T = $1/\alpha$.

This is the basic car-following model that is used in INTRAS but with time-homogeneous processing at 1-sec intervals. Many calculations are needed to form a following trajectory. Although the reaction time T is in the car-following formula, it is not represented well in the simulation.

In EVIPAS, car following is achieved by first setting a target position for the follower in relation to the change of acceleration of the leader and Equation 2. The trajectory of the follower is then calculated from Newton's laws of motion to either pass through the target coordinates or at least a safe position behind the target position. The relative relationship between the current speed and position of the follower and its target speed and position determines the particular set of Newton equations that will be used. Generally the new section of the trajectory can include combinations of acceleration, deceleration, and constant speed.

As the car-following algorithm proceeds, redundant nodes are removed by a filtering process that ensures efficiency in the trajectory fitting. The existence of the complete vehicle trajectory in the simulation means that only those vehicles that affect the controller need to be retained at any time. Generally this includes only the vehicles that have hit the first detector. The simulation therefore is actually dealing with a relatively small number of vehicles at any time.

The vehicle simulation proceeds through a moving window upstream from the stop line, with vehicles processed in order of their position regardless of lane. Green approaches running simultaneously are processed together in the same window. This simplifies the gap-checking procedures in permissive green movements for left turners. The same window format handles a permissive green approach, a protected green approach, a red approach, or an approach with some lanes facing green and some facing red.

Lane changing is allowed in three situations. A through vehicle obstructed by a waiting left-turn vehicle can change into an adjacent through lane, arriving through vehicles will change lanes if a shorter queue is available, and turning vehicles will change into a short turning lane at the start of that lane.

Permissive left-turn vehicles may cross any number of opposing lanes. They are allowed an early start at the beginning of the green and a late turn on yellow. The default values of the probabilities of these maneuvers have been calibrated from field studies but can be changed by the user. Both right- and left-turning vehicles turn at given turning speeds that are derived from the turning radius specified by the user.

There are 100 randomly assigned driver types and driver reaction times; gap-acceptance probability of an early left turn and lane changing are all functions of driver type.

The simulation code is completely structured such that major changes or modifications can be made to one component without affecting other components. The signal controller is currently set up for a one- or two-ring standard NEMA controller. This module could be easily modified to change the existing controller or add a new type controller without affecting the remainder of the model. Similarly the detectors and vehicle actuations have their own module. New detector combinations or actuation procedures, or both, can be easily added without changing the main model.

The intersection simulation has been operating under a wide variety of phasing and detection scenarios. Its real time to computer time ratio is between 1,000 to 1 and 8,000 to 1 for the VAX 8600.

Output Module: OUTPUT

This module summarizes the model outputs including values for the measures of effectiveness and the optimal parameter settings. The overall economic benefit of the improvement is presented.

CALIBRATION AND VALIDATION

Validation of EVIPAS has been undertaken by comparing the simulation with field studies at 10 existing traffic-actuated signalized intersections. Data were collected on traffic volumes, vehicle types, vehicle speeds, stopped delay by approach lane, phase and phase duration, intersection geometry, and timing parameter settings.

All intersections sampled were located across the state of Pennsylvania and included the following types:

1. Fully actuated eight-phase with multilane approaches on the main line.
2. Fully actuated five-phase with volume density and multilane approaches on the main line, with and without volume density.
3. Fully actuated two-phase with permissive left turn.
4. Fully actuated three-phase with a permissive left turn on a multilane approach (i.e., at least two through lanes opposing the permissive left turn).
5. Fully actuated three-phase with leading left turn.
6. Semi-actuated.

For each intersection in the field study, the simulation was calibrated for one data set and then validated by using one or two additional data sets.

Two types of data were compared: the signal timing and the stopped delay of the traffic. The timing comparisons included the average length of each phase and the average cycle length.

TABLE 1 MODEL VALIDATION—SIGNAL TIMING (sec) (INTERSECTION OF ROUTE 19 AND WARRENDALE ROAD, WARRENDALE, PENNSYLVANIA, FIVE-PHASE FULLY ACTUATED)

Data Set	Phase	1+5	2+5	1+6	2+6	3	Cycle
1	Field	13.6	7.9	10.9	26.8	21.0	56.0
	Model	11.2	8.4	7.1	26.1	19.6	55.7
2	Field	13.4	13.7	11.9	24.4	20.8	52.9
	Model	11.2	6.8	9.2	26.2	19.9	56.7
3	Field	14.1	8.5	10.7	29.6	26.1	68.2
	Model	11.2	8.7	7.2	31.5	21.2	64.9

TABLE 2 MODEL VALIDATION—AVERAGE STOPPED DELAY (sec) (INTERSECTION OF ROUTE 19 AND WARRENDALE ROAD, WARRENDALE, PENNSYLVANIA, FIVE-PHASE FULLY ACTUATED)

Data Set	Approach	Lanes	Movement	Field	Model	Volume (vph)
1	SB	2	Through	14	8	450
	SB	1	Left	22	22	59
	WB	1	All	16	12	94
	NB	2	Through	8	5	425
	NB	1	Left	19	19	73
	EB	1	All	22	16	115
2	SB	2	Through	14	7	472
	SB	1	Left	25	26	89
	NB	1	All	14	18	125
	NB	2	Through	8	6	515
	NB	1	Left	28	18	63
	EB	1	All	23	16	82
3	SB	2	Through	11	7	630
	SB	1	Left	32	22	108
	WB	1	All	17	15	150
	NB	2	Through	11	6	762
	NB	1	Left	25	22	67
	EB	1	All	27	29	137

The EVIPAS model replicated the field data very closely. Generally the phase lengths and cycle lengths were within 5 percent of the field results.

The average stopped delay of the traffic was compared for each lane group of each approach. The delay comparisons generally were within 20 percent of the field data. Most cases in which the delays did not agree very well could be traced to irregular detector performance, or local peculiarities in driver behavior with regard to the observance of lane directions.

Tables 1 through 8 give examples of the comparisons between field observations and the computer model.

CONCLUSION

The EVIPAS model is showing considerable promise for the evaluation and optimization of a variety of types of traffic signal installations. The development efforts have concentrated on producing a general capability to model most geometric, traffic, and control scenarios and to provide an efficient and rigorous optimization structure.

The model has been programmed to allow future changes in the controller or detection, or both, without any modifications to the main program. Testing and validation of EVIPAS has shown that it replicates observed vehicle behavior and controller phasing.

ACKNOWLEDGMENTS

The research project on which this paper is based was sponsored by the Pennsylvania Department of Transportation and the Federal Highway Administration, U.S. Department of Transportation, through the Pennsylvania Department of Transportation Research Project RPN-84-10, "Enhancement of VIPAS."

The original VIPAS model was programmed by John Breon of the Pennsylvania Department of Transportation. Personnel from the Universitv of Pittsburgh who have assisted in the programming and field studies include Mohamed Belker, Adanan Hussein, Mohamed Mohamed, Kim Moio, Dimitrios Papaikolaou, and Linda Zemotel.

TABLE 3 MODEL VALIDATION—SIGNAL TIMING (sec) (INTERSECTION OF ROUTE 212 AND PERITAN STREET, UNIONTOWN, PENNSYLVANIA, SEMI-ACTUATED)

Data Set	Phase	1	2	Cycle
1	Field	19.4	84.8	104.2
	Model	18.8	82.8	101.0
2	Field	20.0	81.9	101.9
	Model	19.0	81.5	99.9

TABLE 4 MODEL VALIDATION—AVERAGE STOPPED DELAY (sec) (INTERSECTION OF ROUTE 212 AND PERITAN STREET, UNIONTOWN, PENNSYLVANIA, SEMI-ACTIVED)

Data Set	Approach	Lanes	Movement	Field	Model	Volume (vph)
1	SB	1	All	17	17	38
	WB	1	All	1	2	173
	NB	1	All	17	22	19
	EB	1	All	3	2	231
2	SB	1	All	19	19	42
	WB	1	All	4	2	223
	NB	1	All	19	9	26
	EB	1	All	3	2	307

Note: EB, WB is the main highway and is not actuated.

TABLE 5 MODEL VALIDATION—SIGNAL TIMING (sec) (INTERSECTION OF ROUTE 30 AND ROUTE 48, IRWIN, PENNSYLVANIA, EIGHT-PHASE FULLY ACTUATED)

Data Set	Phase	1+5	2+5	1+6	2+6	3+7	4+7	3+8	4+8	Cycle
1	Field	23.7	None	14.7	50.6	23.0	7.5	2.8	35.2	145.3
	Model	21.7	3.5	13.9	51.0	19.8	10.1	6.8	38.2	145.9
2	Field	23.1	None	13.4	47.2	20.4	8.0	4.5	36.6	139.6
	Model	21.1	4.1	13.1	54.5	20.5	16.3	3.6	38.8	153.6

TABLE 6 MODEL VALIDATION—AVERAGE STOPPED DELAY (sec) (INTERSECTION OF ROUTE 30 AND ROUTE 48, IRWIN, PENNSYLVANIA, EIGHT-PHASE FULLY ACTUATED)

Data Set	Approach	Lanes	Movement	Field	Model	Volume (vph)
1	SB	2	All	62	65	403
	WB	3	All	39	37	642
	NB	2	All	NA	48	384
	EB	3	All	33	31	651
2	SB	2	All	65	69	566
	WB	3	All	43	40	596
	NB	2	All	NA	56	401
	EB	3	All	52	48	1105

TABLE 7 MODEL VALIDATION—SIGNAL TIMING (sec) (INTERSECTION OF ROUTE 322 AND CHURCH STREET, STATE COLLEGE, PENNSYLVANIA, TWO-PHASE FULLY ACTUATED)

Data Set	Phase	1	2	Cycle
1	Field	43.2	13.4	56.6
	Model	42.9	15.9	58.2
2	Field	41.9	15.7	57.6
	Model	42.3	15.8	57.5

TABLE 8 MODEL VALIDATION—AVERAGE STOPPED DELAY (sec) (INTERSECTION OF ROUTE 322 AND CHURCH STREET, STATE COLLEGE, PENNSYLVANIA, TWO-PHASE FULLY ACTUATED

Data Set	Approach	Lanes	Movement	Field	Model	Volume
1	SB	2	All	3	3	984
	WB	1	All	21	17	62
	NB	2	All	4	3	388
	EB	1	All	17	16	132
2	SB	2	All	5	3	1097
	WB	1	All	19	18	108
	NB	2	All	5	2	452
	EB	1	All	18	29	115

REFERENCES

1. *Special Report 209: Highway Capacity Manual.* TRB, National Research Council, Washington, D.C., 1985.
2. *SOAP84, User's Manual.* Report FHWA-IP-85-8. FHWA, U.S. Department of Transportation, 1985.
3. F. V. Webster. *Traffic Signal Settings.* Road Research Paper 39, Department of Scientific and Industrial Research, HMSO, London, England, 1958.
4. T. W. Rioux and C. E. Lee. Texas—Microscopic Traffic Simulation Package for Isolated Intersection. Presented at the 56th Annual Meeting of the Transportation Research Board, Washington, D.C., 1977.
5. *Traffic Network Analysis with NETSIM—User's Guide.* Report FHWA-IP-80-3. FHWA, U.S. Department of Transportation, 1980.
6. *Value Iteration Process—Actuated Signals VIPAS.* Vol. 1. Model Design and Components. Pennsylvania Department of Transportation, Harrisburg, Pa., 1983.
7. *Development and Testing of INTRAS, A Microscopic Freeway Simulation Model.* Vol. 2. Report FHWA/RD-80-107. FHWA, U.S. Department of Transportation, 1980.
8. E. W. Wright. The Use of Conversational Computer Programs in the Structural Design Engineering Office. *Canadian Journal of Civil Engineering,* Vol. 4, No. 4, 1977, pp. 417–435.
9. J. H. Kell. Analyzing Vehicular Delay at Intersections Through Simulation. *Bulletin 356,* HRB, National Research Council, Washington, D.C., 1962, pp. 28–39.
10. D. I. Robertson. *Transit: A Traffic Network Study Tool.* Road Research Laboratory Report RL-253. Crowthorne, England, 1969.
11. J. E. Dennis and R. B. Schnabel. *Numerical Methods for Unconstrained Optimization and Non-Linear Equations.* Prentice-Hall Book Co., Englewood Cliffs, N.J., 1983.

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 either the Federal Highway Administration, U.S. Department of Transportation, or the Commonwealth of Pennsylvania. The paper does not constitute a standard, specification, or regulation.

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

Interlibrary Loan Request Form

Request No.: Date: 03/04/03 Need before: 03/06/03 Notes: 2013589-1110
Call No.:

McGuireWoods LLP
Attention: Library
One James Center
901 E. Cary Street
Richmond, Virginia 23219

Patron information: Sam Tarry

Book author, OR, Serial title, volume, issue, date, pages; OR Audiovisual title:
Transportation Research Record, 1114, 1987, 111-22.

Book title, edition, imprint series; OR, Article author, title: This edition only
Bowman, B. L., The Effectiveness of Railroad Constant Warning Time Systems.

Verified in; AND/OR Cited in:
ISBN, ISSN, LCCN, or other bibliographic number:

Transportation Research Board Library
500 Fifth St. NW
Washington, DC 20001

Request complies with
[ ] 108(g) (2) Guidelines (CCG)
[ ] other provision of copyright law (CCL)

Authorization: Doris Morgan
Telephone: 804.775.7618

TYPE OF REQUEST:
[ ] LOAN; WILL PAY FEE ______
[X] PHOTOCOPY; MAX COSTS $20
[ ]

LENDING LIBRARY REPORT: Date
Date shipped: ____________ via ______
Insured for $ ________ Charge $ _____
DUE ____________ [ ] Return insured
Packing Requirements

RESTRICTIONS: [ ] Library use only
[ ] Copying not permitted
[ ] No renewals [ ] ____________

NOT SENT BECAUSE:
[ ] In use [ ] Lacking [ ] Not owned
[ ] At bindery [ ] Cost exceeds limit
[ ] Non Circulating [ ] Lost
[ ] Not found as cited [ ] On order
[ ] Not found on shelf
[ ] Lacks copyright compliance
[ ] In process [ ] Request on _____
[ ] Hold placed [ ] Poor Condition
[ ] Estimate Cost of Loan $ ________
Photocopy $ ____ Microfilm/fiche $ ____
[ ] Prepayment required

BORROWING LIBRARY REPORT:
Date received _____ Date returned _____
Returned via ________ Insured for $ ____
Payment provided $ ____________

RENEWALS:
Date requested ____________
New due date ____________
Renewal denied ____________

The Effectiveness of Railroad Constant Warning Time Systems

BRIAN L. BOWMAN

Presented in this paper are the results of two tasks of a study sponsored by the Federal Highway Administration. The purpose of these tasks was to determine the effectiveness of railroad constant warning time (CWT) systems in (a) reducing motorists violation of activated at-grade warning systems, and (b) reducing vehicle-train accidents. CWT systems have the capability of measuring train motion, direction of movement, and distance from the crossing. These parameters are interpreted by the control logic to provide estimates of train speed and arrival time. When the estimated arrival time achieves a preselected minimum, such as 20 sec, the warning displays at the crossing are activated. Analysis of operational data indicated that CWT systems are effective in providing both a uniform amount of advance warning and in reducing motorist violation of the warning system. A comparative analysis of vehicle-train accidents occurring from 1980 through 1984 was also performed. This analysis indicated that, in the majority cases, crossings with CWT systems have a lower accident rate than crossings without CWT. Nevertheless, this difference was not large enough to be statistically significant at the 95 percent confidence level.

The ability to command the respect of motorists is a key factor in establishing the effectiveness of traffic control devices. A genuine need, proper device placement, and consistent operation are all important in obtaining and retaining motorist respect. Failure to consider these factors leads to motorist contempt, disregard for traffic controls, and potentially to accidents.

Train-activated traffic controls at railroad-highway grade crossings are particularly susceptible to the loss of motorists' respect. This is primarily the result of variations in warning time and the need for fail-safe design. The majority of train-activated devices now in use are based on track circuits and control logic initially developed approximately 100 years ago. This system is based on an approach track circuit length designed to provide a preselected warning time for the fastest train. The use of island circuits permits the system to determine train direction and cease signal operation after the train has passed the crossing. Such a system, unless configured with overriding capabilities, provides continuous detection while a train is on the approach. Trains traveling slower than the design speed or stopping on the approach length result in prolonged activation of the railroad-highway warning system.

The fail-safe design is required because the crossing warning devices are active in the presence of a train and unactivated at all other times. The absence of the flashing lights is intended to indicate to the motorist that it is safe to proceed. This requires that the warning system be provided with standby power in case of a commercial power failure and that the system revert to the active mode if failure of an element or component of the system, including the rails, occurs. Prolonged and fail-safe activation have resulted in motorists often disregarding the warning and driving through or around the warning devices *(1)*. Accident statistics indicate that more than 49 percent of all train-involved accidents and 45 percent of crossing fatalities occur at locations with some form of active warning *(2)*.

The potential consequences associated with excessively long warning times resulted in the development of a constant warning time (CWT) track circuit and control logic system. The CWT system, developed during the 1960s, differs from other systems in that it is capable of detecting train speed in addition to train motion, direction, and distance from the crossing. The ability to measure train speed and distance from the crossing enables a continuous update on the actual arrival time. When the estimated arrival time achieves a preselected minimum, such as 20 sec, the warning displays at the crossing are activated. Trains that enter the approach section and subsequently stop or reverse direction without reaching the roadway crossing are interpreted by the control logic as not requiring activation of the crossing warning system. Motorists are not, therefore, subjected to long delays caused by slow or stopped trains and can expect the arrival of a train within a uniform and reasonable length of time following the initiation of the crossing controls.

The research reported here was sponsored by the FHWA to determine how effective CWT systems are in reducing vehicle-train accidents and increasing motorist compliance with activated at-grade warning devices. This task was accomplished by analyzing data obtained from railroads, individual states, the Federal Railroad Administration, and operational data collected at railroad crossings.

EFFECTIVENESS OF CONSTANT WARNING TIME SYSTEMS IN REDUCING ACCIDENTS

The selection of accident-based measures of effectiveness was based on the probable impact of providing a uniform amount of warning time. This involved analyzing only those accidents where the roadway vehicle was struck by or strikes the first unit of the train. The rationale behind this analysis was that motorists who believe that there is an excessive amount of warning time will cross in front of an oncoming train after stopping or try to race the train to the crossing. Accidents where the train was fully in the crossing and the roadway vehicle strikes subsequent train units cannot be corrected by the installation of CWT systems. These accidents are more a result of driver

Goodell-Grivas, Inc., 17320 W. Eight Mile Road, Southfield, Mich. 48075.

inattention, excessive speed, sight restrictions, or improper warning device operation than the influence of train detection and control logic systems used at the crossing.

Accidents where the train struck the vehicle and the vehicle struck the first unit of the train were further stratified into two categories: (a) characteristics of the accident and (b) physical and operational characteristics of the crossing.

Site Selection Criteria

The effectiveness of CWT systems in reducing accidents was determined by performing analyses between different combinations of warning devices and track circuit-control logic systems. The following combinations of crossing types were used in the analysis:

- Flashing lights without CWT,
- Flashing lights with CWT,
- Gates without CWT, and
- Gates with CWT.

The site selection process was initiated by stratifying the FRA's national inventory, by crossing type, into categories of average daily traffic (ADT) and trains per day. Approximately 60 crossings, for each device type, were randomly selected from the cells that maximized ADT and train volumes. The complete inventory for each crossing was obtained and the operating railroad and the geographic location of the crossing were identified. Information was requested from the railroads to verify the type of warning device and track circuit and the respective date of installation as well as operational and physical characteristics of the crossing. When possible, the respective highway agencies were also contacted to request updates on the number of roadway lanes and ADT counts. If verified information pertaining to the type of warning device and the presence of a CWT system was not received on a crossing, it was eliminated from further analysis. A flowchart of the site selection and verification process is shown in Figure 1.

The number of crossings verified for each crossing type, and subsequently used in the accident analysis, is summarized in Table 1. The smallest number of crossings occurs in the flashing light with CWT category because there are relatively few crossings that have flashing lights with CWT capabilities. The majority of CWT installations occur in conjunction with gates.

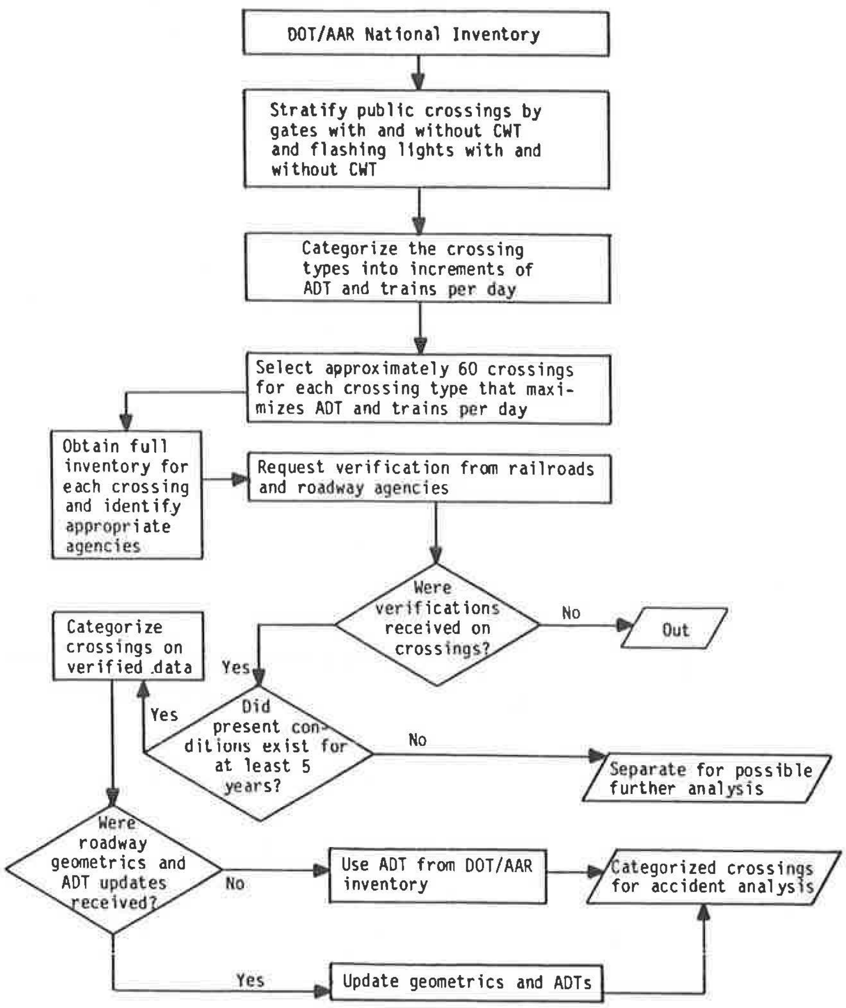


FIGURE 1 Flowchart of site selection and verification process used for accident analysis.

TABLE 1 NUMBER OF CROSSINGS WITH VERIFIED TYPES OF WARNING AND TRACK CIRCUITRY DEVICES USED FOR ACCIDENT ANALYSIS

	Gates With CWT	Gates Without CWT	Flashing Lights With CWT	Flashing Lights Without CWT
Number of crossings	27	39	13	26

Many of the replies returned for flashing lights with CWT indicated that either CWT systems were not in place or gates had been installed.

Measure of Exposure

Comparative accident analysis between independent groups requires the use of exposure rates because the probability of an accident occurring is directly related to the number of available opportunities. For train-involved crossing accidents, the number of opportunities are represented by the roadway volume and the amount of time that the crossing is occupied by the train. The only exposure factors that are prominent in analyzing the effectiveness of CWT installations are, however, roadway and train volumes. This is because the only accidents that can be reasonably associated with the effect of CWT systems are those occurring with the first unit of the train. Determination of train occupancy time at the crossing is, therefore, not required. The exposure measure used in the analysis to obtain the accident rate is displayed as follows:

$$\text{Accident rate} = \frac{(\text{number of accidents})\ (1 \times 10^9)}{(\text{ADT})\ (\text{trains per day})\ (365)\ (\text{years})}$$

Results of Accident Analysis

A search of the computerized train-involved accident files provided by the FRA was performed for all of the crossings that were verified as possessing the required warning and track circuitry devices. Information pertaining to crossing geometrics, operational data, and accident characteristics were coded for computer analyses. Analyses were performed on all accidents occurring from 1980 through 1984.

Summaries of accident frequency categorized by accident characteristics and physical-operational characteristics are given in Tables 2 through 4. Because a different number of

TABLE 2 SUMMARY OF ACCIDENT TYPES FOR YEARS 1980 TO 1984

	Crossing Type			
Accident Type	Gates With CWT	Gates Without CWT	Flashing Lights With CWT	Flashing Lights Without CWT
Struck by train	8	16	5	17
Striking first unit of train	0	1	2	4
Striking other unit of train	2	1	3	0
Total	10	18	10	21

TABLE 3 SUMMARY OF ACCIDENT CHARACTERISTICS REPRESENTED AS FREQUENCIES

Accident Characteristics	Gates With CWT		Gates Without CWT		Flashing Lights With CWT		Flashing Lights Without CWT	
	Struck	Striking 1st Unit	Struck	Striking 1st Unit	Struck	Striking 1st Unit	Struck	Striking 1st Unit
Driver Action								
Drove around or through	2	0	5	0	0	0	0	0
Stopped and then proceeded	1	0	1	0	0	0	3	0
Did not stop	2	0	4	1	5	2	8	2
Other	2	0	4	0	5	0	8	1
Unknown	3	0	6	0	0	0	6	1
Severity								
Fatal	0	0	2	0	0	0	2	0
Personal injury	2	0	2	0	0	0	8	2
Property Damage only	6	0	12	1	5	2	7	2

TABLE 4 SUMMARY OF ACCIDENT FREQUENCY CATEGORIZED BY PHYSICAL AND OPERATIONAL CHARACTERISTICS PRESENT AT TIME OF ACCIDENTS

Physical or Operational Characteristics	Gates With CWT		Gates Without CWT		Flashing Lights With CWT		Flashing Lights Without CWT	
	Struck	Striking 1st Unit	Struck	Striking 1st Unit	Struck	Striking 1st Unit	Struck	Striking 1st Unit
Crossing Angle								
0-29	0	0	0	0	0	0	0	0
30-60	1	0	1	0	1	0	1	0
60-90	7	0	15	1	4	2	16	4
Number of Tracks								
1	1	0	4	0	2	0	10	0
2	3	0	10	0	2	0	5	3
3	2	0	0	0	0	0	2	1
>3	2	0	2	1	1	2	0	0
Maximum train speed (mph)								
<10	0	0	0	0	3	2	0	0
11-20	1	0	3	0	0	0	0	1
21-40	3	0	3	0	1	0	4	1
41-60	4	0	4	1	1	0	8	0
>60	0	0	6	0	0	0	5	2
Train Speed Ratio								
<2:1	3	0	6	0	1	0	9	3
2:1	2	0	0	0	2	1	0	0
3:1	0	0	1	0	0	0	2	1
>3:1	3	0	9	1	2	1	6	0
Switching Ratio								
0	1	0	6	0	2	0	9	3
0.1-0.9	1	0	0	0	0	0	0	1
1.0-1.9	1	0	0	0	0	0	0	0
2.0-2.9	5	0	3	0	1	1	6	0
3.0-3.9	0	0	3	1	1	0	2	0
4.0-5.9	0	0	0	0	1	1	0	0
6.0-7.9	0	0	1	0	0	0	0	0
>8.0	0	0	3	0	0	0	0	0

crossings with indigenous ADT and train volumes comprise the population of each crossing category, it was necessary to normalize the accident frequencies by the 5-year exposure. The exposure measure used for accident type and accident characteristics was based on the total 5-year exposure for each crossing type as presented in Table 5.

For the purposes of analysis it was necessary to combine these categories that had no crossings with the attributes being analyzed with adjacent categories to reduce the number of missing values. When feasible, those instances of zero accidents were also combined with adjacent categories. When this occurred, the exposure rate of the adjacent categories was also used in determining the accident rate. A summary of the accident frequency for the physical and operational characteristics is given in Table 6.

The data were analyzed by performing the Mann-Whitney *U*-test on the accident rates. The rates were determined by adding accidents in which the vehicle was struck by the train and struck the first unit of the train. This sum was then divided by the appropriate measure of exposure. This nonparametric test was used to determine if the independent categories of similar warning devices with and without CWT were from the same population. All of the tests were conducted at a 95 percent level of confidence. If the two-tailed probability of

TABLE 5 FIVE-YEAR TOTAL ACCIDENT EXPOSURE FACTOR (BILLION VEHICLE TRAINS) AND NUMBER OF CROSSINGS IN EACH CATEGORY

Crossing Type							
Gates With CWT		Gates Without CWT		Flashing lights with CWT		Flashing lights without CWT	
Number	Exposure	Number	Exposure	Number	Exposure	Number	Exposure
27	12.40	39	14.00	13	4.39	26	8.83

TABLE 6 SUMMARY OF THE NUMBER OF CROSSINGS AND THE 5-YEAR EXPOSURE (BILLION VEHICLE TRAINS) FOR SELECTED PHYSICAL AND OPERATIONAL CROSSING CHARACTERISTICS

Crossing Characteristics	Gates with CWT		Gates without CWT		Flashing lights with CWT		Flashing lights without CWT	
	Number	Exposure	Number	Exposure	Number	Exposure	Number	Exposure
Crossing Angle								
0-29	2	0.77	1	0.21	0	0	0	0
30-60	4	1.14	4	1.85	3	0.87	3	0.56
61-90	21	10.50	34	11.90	10	3.52	23	8.27
Number of Tracks								
1	11	4.09	5	1.72	9	3.07	8	2.57
2	6	4.20	22	8.30	3	1.00	8	2.62
3	3	1.32	9	2.99	0	0	6	2.16
>3	7	2.77	3	0.95	1	0.32	4	1.49
Maximum train speed (mph)								
<10	0	0	2	0.73	2	0.72	6	2.37
11-20	4	1.18	7	2.28	1	0.43	6	1.52
21-40	9	3.94	12	4.85	4	1.29	7	2.77
41-60	11	5.81	10	3.16	5	1.74	3	1.37
>60	3	1.45	8	2.94	1	0.22	4	0.80
Train Speed Ratio								
<2:1	1	0.57	15	5.88	6	2.22	6	1.80
2:1	4	2.72	2	0.76	1	0.34	2	0.69
3:1	0	0	6	1.97	1	0.22	10	3.33
>3:1	22	9.10	16	5.65	5	1.60	8	3.01
Switching Ratio								
0	4	2.12	10	3.64	5	1.82	11	3.47
0.1-0.9	4	1.97	3	0.81	0	0	6	1.79
1.0-1.9	6	1.47	4	1.01	3	1.02	1	0.46
2.0-2.9	6	3.47	5	1.87	0	0	2	0.96
3.0-3.9	3	1.34	6	2.02	1	0.20	2	0.95
4.0-5.9	0	0	2	0.99	2	0.50	1	0.39
6.0-7.9	3	1.38	4	1.64	0	0	0	0
>8.0	1	0.65	5	1.98	2	0.85	3	0.81

occurrence from the test was equal to or less than 5 percent, it was concluded that CWT systems had an impact on accidents.

Inspection of Tables 7 through 9 indicates that there were no significant differences at the 95 percent confidence level in the distribution of accident rates between crossings with CWT systems and those without. The accident rate at crossings equipped with CWT systems was in the majority of instances lower than comparable crossings without CWT systems. This difference was not large enough, however, to state with a 95 percent level of confidence that accident rates are lower at crossings equipped with CWT systems.

COLLECTION AND ANALYSIS OF OPERATIONAL DATA

Traffic accidents are the most acceptable and widely used measure of highway safety. However, the stochastic nature of accidents requires relatively large sample sizes collected over long periods of time. This does not pose a problem for locations with high accident frequencies but for relatively low-accident frequency locations, such as at-grade railroad crossings, the use of accident statistics becomes increasingly problematic. As a result of the recognized shortcomings associated

TABLE 7 RESULTS OF MANN-WHITNEY *U*-TEST ON THE ACCIDENT RATES (ACCIDENTS PER BILLION VEHICLE TRAINS) FOR ACCIDENT TYPE

Accident Type	Crossing Type: Gates with CWT	Gates without CWT	Flashing lights with CWT	Flashing lights without CWT
Struck by Train	0.645	1.143	1.139	1.925
Striking 1st unit	0	0.071	0.456	0.453
Striking other unit	0.161	0.071	0.683	0
Test statistic and 2-tail probability	Z = 0.2214 P = 0.8248		Z = 0.6457 P = 0.5127	

TABLE 8 RESULTS OF MANN-WHITNEY *U*-TEST ON THE ACCIDENT RATES (ACCIDENTS PER BILLION VEHICLE TRAINS) FOR CHARACTERISTICS OF THE ACCIDENT

Accident Characteristics	Crossing Type: Gates with CWT	Gates without CWT	Flashing lights with CWT	Flashing lights without CWT
Driver Action				
Drove around or through	0.161	0.357	0	0
Stopped and then proceeded	0.081	0.071	0	0.340
Did not stop	0.161	0.357	1.595	1.133
Other	0.161	0.286	1.139	1.019
Unknown	0.242	0.429	0	0.793
Test statistic and 2-tail probability	Z = 1.5910 P = 0.1116		Z = 0.2155 P = 0.8294	
Severity				
Fatal	0	0.143	0	0.227
Personal injury	0.161	0.143	0	1.133
Property Damage only	0.484	0.929	1.595	1.019
Test statistic and 2-tail probability	Z = 0.2214 P = 0.8248		Z = 0.6642 P = 0.5066	

TABLE 9 RESULTS OF MANN-WHITNEY *U*-TEST ON THE ACCIDENT RATES (ACCIDENTS PER BILLION VEHICLE TRAINS) FOR PHYSICAL AND OPERATIONAL CHARACTERISTICS OF THE CROSSING

Physical and Operational Characteristics	Crossing Type: Gates with CWT	Gates without CWT	Flashing lights with CWT	Flashing lights without CWT
Crossing Angle				
0-60	0.523	0.485	0.115	0.180
61-90	0.667	1.345	1.705	2.418
Test statistic and 2-tail probability	Z = 0.0000 P = 1.0000		Z = 0.7746 P = 0.4386	
Number of Tracks				
1	0.244	2.326	0.651	3.891
2	0.714	1.205	2.002	3.053
>3	0.978	0.762	9.434	0.822
Test statistic and 2-tail probability	Z = 1.5275 P = 0.1266		Z = 0.2182 P = 0.8273	
Maximum train speed (mph)				
0-19	0.847	1.316	4.348	0.257
20-39	0.761	0.619	0.775	1.805
>40	0.551	1.803	0.512	6.912
Test statistic and 2-tail probability	Z = 1.0911 P = 0.2752		Z = 0.2182 P = 0.8273	
Train Speed Ratio				
<2:1	5.245	1.075	0.450	6.667
2:1,3:1	0.735	0.366	5.319	0.747
>3:1	0.330	1.770	1.875	1.993
Test statistic and 2-tail probability	Z = 0.2182 P = 0.8273		Z = 0.6547 P = 0.5127	
Switching Ratio				
0.0-0.9	0.489	1.349	1.099	2.472
1.0-2.9	1.214	1.042	0.980	4.222
>3	0	1.206	1.931	0.929
Test statistic and 2-tail probability	Z = 1.0911 P = 0.2752		Z = 0.6547 P = 0.5127	

with using accidents as the sole measure of safety, the accident analysis was complemented with observations of driver behavior. This analysis was conducted at 12 railroad crossings with the following CWT-crossing control combinations:

- Three crossings with automatic gates and CWT systems.
- Three crossings with automatic gates and no CWT systems.
- Three crossings with flashing lights (only) and CWT systems.
- Three crossings with flashing lights (only) and no CWT systems.

Selection of Measures of Effectiveness

Constant warning time systems are intended to have an indirect impact on accidents by increasing the credibility of at-grade warning devices. This increase in credibility results from the ability of CWT systems to provide a uniform amount of warning time until train arrival at the crossing. The uniform warning time is intended to provide motorists with a consistent expectation of train arrival thereby resulting in fewer violations of the flashing lights and, hence, fewer train accidents. The relationship between the intended purpose of CWT systems, the intermediate objectives, and the ultimate objective of reducing accidents is shown in the causal chain of Figure 2.

The collection of field data was concentrated on obtaining quantifiable measures of effectiveness that (a) indicated whether CWT systems actually do provide a uniform warning time and that (b) could be directly related to the intermediate objectives. The measures of effectiveness selected for the study are given in Table 10.

Test Site Selection Procedure

The measures of effectiveness determined as being appropriate for the analysis of the operational CWT data required observations of motorists' action only during the activated state. In addition, the observational opportunities during the activated state, in most instances, were only present for the first vehicle on each approach lane. This necessitated that the site selection process consider only those crossings with relatively high vehicle and train volumes to maximize the observational opportunities. Other key locational characteristics were desired to help ensure homogeneity between analysis sites. This homogeneity was necessary to increase the probability that observed differences between the test sites were a result of the train detection and type of warning device and not extraneous factors. The key locational variables for which similarities between the 12 locations were desired included

- Sight distance to crossing flashers on the approach,
- Number of tracks,
- Railroad-highway intersecting angle,
- Sight distance along the tracks,
- Roadway grade, and
- Elevation of railroad-highway crossing with respect to roadway elevation.

The initial site selection process was performed by selecting crossings that had been verified as having CWT systems for the accident analysis. Each prospective site was visited to determine the presence of a suitable observer refuge area, proper warning device, and correct locational variables. The respective highway agencies and operating railroads were then contacted for those sites that satisfied all of the preliminary selec-

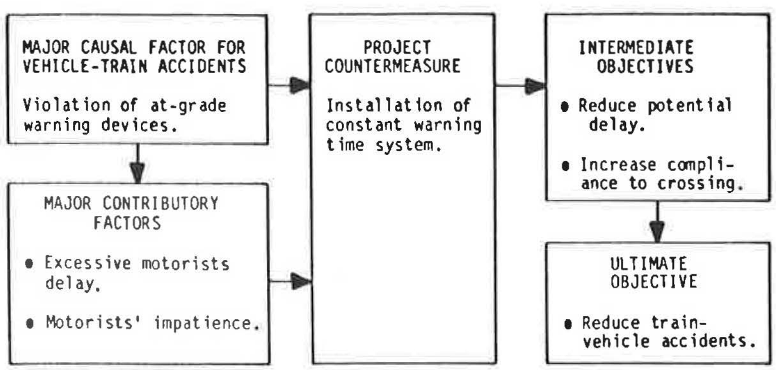


FIGURE 2 Causal chain for the reduction of vehicle-train accidents by installing CWT systems.

TABLE 10 RELATIONSHIP OF MEASURES OF EFFECTIVENESS TO ANALYSIS OBJECTIVES

Purpose	Measure of Effectiveness
To determine if CWT systems provide a uniform amount of warning time.	Warning time until train arrival analyzed in conjunction with train speed.
To determine if CWT systems reduce vehicle delay.	Warning time until train arrival.
To determine if CWT systems result in increased vehicle compliance to warning devices.	Violation rate.

tion criteria. These contacts provided information pertaining to hourly roadway counts, daily train volume, train schedule, and additional verification of the type of train detection and control logic present at the site. Twelve locations, three in each category of train detection system and warning device, were selected that maximized train and vehicle exposure.

Field Data Collection Procedure

Data were obtained manually with the use of radar guns and stop-watches. One observer was placed on each crossing approach. The stopwatches were initiated on first activation of the warning device at the crossing. The observers noted the time of vehicle arrival for the first vehicle in each lane, the time of violation if the flashers were activated, the time of train arrival and departure, and the speed of the train. Violation time was recorded for each vehicle that went through the activated flashers or that drove around the gates. The time of arrival for each vehicle that had the opportunity to violate (the first vehicle in the queue of each lane) was the time at which the vehicle arrived at the stop-bar of the approach.

Analysis of Operational Data

Effectiveness of CWT in Providing Uniform Warning Time

The variations in train speed given in Table 11 indicate that accompanying variations in warning time could be expected at each crossing. This variation in warning time would be proportional to the train speed unless the train detection and control logic compensated for the variation. For example, for crossings without CWT capabilities, if 30 sec was the observed warning time at 40 mph (64 km/h), then 240 sec (8 times 30 sec) would be required for a train traveling 5 mph (8 km/h). The track circuits and control logic prevented this wide variation in warning time from occurring at all of the crossings studied. Those crossings that were not equipped with CWT systems were equipped with motion sensors. The observed instances of very low speeds were caused by switching activities in the approach circuit before the train entered the crossing. Therefore, the lower train speeds were the result of trains accelerating from a stop on the approach circuit.

The effectiveness of CWT systems in providing uniform warning times was analyzed by performing an analysis of variance (ANOVA) and plotting intervals of train speed versus average warning time. The results of the two-way analysis of variance given in Table 12 indicate that there is a significant difference at the 95 percent level of confidence between the effect of the different types of crossings and the average warning times. This difference was further analyzed with the Scheffe contrast test to determine where these differences resided. The results of the Scheffe test given in Table 13 indicate that there are significant differences, at a 95 percent level of confidence, between crossings equipped with CWT systems and those without such systems. Crossings equipped with CWT systems, therefore, display different characteristics in their

TABLE 11 MAXIMUM, MINIMUM, AND STANDARD DEVIATION OF TRAIN VELOCITIES (mph) OBSERVED BY TYPE OF CROSSING (1 mph = 1.6 km/h)

Parameter	Flashing lights without CWT	Flashing lights with CWT	Gates without CWT	Gates with CWT
Maximum speed	41	31	44	35
Minimum speed	5	1	3	2
Standard deviation	9.3	17.5	17.0	12.9
Ratio of minimum to maximum speed	1:8	1:31	1:15	1:18

TABLE 12 ANOVA ON THE MEAN WARNING TIME (sec) PER TRAIN VELOCITY GROUP (mph) FOR DIFFERENT CROSSING TYPES

Speed Group	Crossing Type: Flashing lights without CWT	Flashing lights with CWT[1/]	Gates without CWT	Gates with CWT[1/]
0-5	81.6	35.5	57.5	36.3
6-10	77.6	35.0	47.8	32.2
11-15	80.6	27.0	49.5	31.7
16-20	68.8	30.8	65.2	33.0
21-25	60.4	30.1	68.6	33.0
26-30	50.3	34.4	50.1	37.2
31-35	43.2	33.0	50.5	29.2
36-40	33.0	19.9	40.0	38.0
>40	48.9	33.0	42.0	38.0

Source	df	SS	MS	F_{ij}	95% critical F value
Crossing type	8	3535.2	441.9	2.43*	2.38
Speed group	3	1251.3	417.08	2.29	3.03
Error	23	4190.0	182.17		

1/ - missing value estimated to minimize SS error
1 mi/h = 1.6 km/h
Asterisk (*) indicates significance

TABLE 13 SCHEFFE CONTRAST TEST ON THE EFFECT OF CROSSING TYPE ON MEAN WARNING TIME (sec)

	Flashing lights without CWT	Flashing lights with CWT	Gates without CWT	Gates with CWT
Flashing lights without CWT	----	----	----	----
Flashing lights with CWT	265.7*	----	----	----
Gates without CWT	73.2	192.5*	----	----
Gates with CWT	235.8*	29.9	162.6*	----

95 percent Scheffe contrast value = 159.3
1 mi/h = 1.6 km/h
Asterisk (*) indicates significant difference.

average warning time than crossings not equipped with CWT systems.

The values in Table 12 were plotted and the linear best fit regression line obtained. An inspection of these plots, shown in Figure 3, indicates a negative slope for all crossing types except for gates with CWT. With one exception, this indicates that as train velocity increases, the amount of advance warning time decreases. The linear approximation for crossings with flashing lights and CWT has the least slope. The presence of a truly uniform warning time would be characterized by a slope of zero magnitude. Because crossings with CWT are closer to the desirable zero slope, the differences demonstrated by the ANOVA and Scheffe contrast tests can be interpreted as differences in uniformity of warning time. Crossings equipped with CWT systems do, therefore, provide a more uniform warning time to motorists.

Effectiveness of CWT in Reducing Warning Time Violation

Each of the crossings at which data were collected was located on relatively high-volume roadways. The high volumes resulted in a queue of vehicles on each approach lane at every test crossing during activation of the warning devices. The occupied roadway approaches resulted in the number of opportunities for vehicles to proceed through the activated warning devices (violations) being similar, per unit of time, for each test site. Because the violation opportunities are time dependent, however, a greater number of opportunities exist when the amount of time from device activation to train arrival is increased.

The effectiveness of CWT systems in reducing violations of the warning system was determined by analyzing violations in conjunction with both the total amount of warning time and the

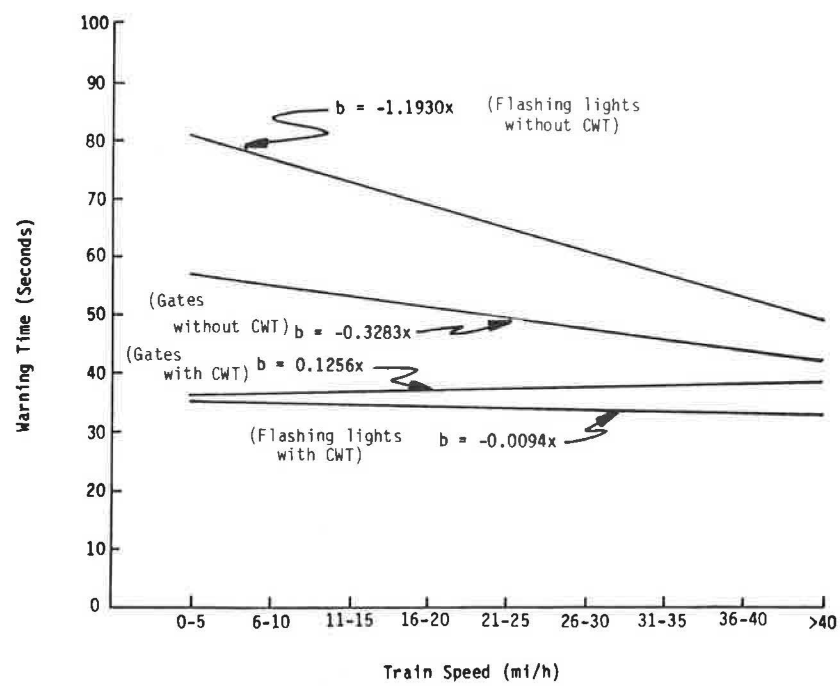


FIGURE 3 Best fit linear approximations and the resultant slopes for each crossing type on speed groups and mean warning time.

TABLE 14 OBSERVED VIOLATIONS OF THE ACTIVATED WARNING DEVICE CATEGORIZED BY TOTAL WARNING TIME FOR DIFFERENT CROSSING TYPES

Total Warning Time (Seconds)	Number of Violations by Crossing Type: Flashing lights without CWT	Flashing lights with CWT	Gates without CWT	Gates with CWT
11-15	0	0	2	0
16-20	0	0	1	0
21-25	3	0	1	0
26-30	7	33	5	2
31-35	6	30	1	14
36-40	25	27	2	4
41-45	41	4	4	0
46-50	22	0	9	0
>50	265	0	192	0
Totals	369	94	217	20

TABLE 15 OBSERVED VIOLATION OF THE ACTIVATED WARNING DEVICE AND CUMULATIVE PROPORTIONS CATEGORIZED BY TIME UNITL TRAIN ARRIVAL FOR DIFFERENT CROSSING TYPES

Time until train arrival (seconds)	Number of Violations by Crossing Type: Flashing lights without CWT	Flashing lights with CWT	Gates without CWT	Gates with CWT
0-5	1	1	0	3
6-10	17	4	3	2
11-15	34	13	13	4
16-20	30	26	13	4
21-25	35	20	18	6
26-30	38	19	17	1
31-35	29	10	11	0
36-40	29	1	20	0
>40	156	0	122	0
Totals	369	94	217	20

TABLE 16 KOLMOGOROV-SMIRNOV TEST ON THE NUMBER OF VIOLATIONS OCCURRING WITHIN CATEGORIES OF ADVANCE WARNING TIME (sec) FOR CROSSINGS EQUIPPED WITH GATES

Total Warning Time Interval	Gates Without CWT		Gates With CWT		Absolute Differences in Cumulative Occurrences
	Occurrences	Cumulative Occurrences	Occurrences	Cumulative Occurrences	
0-5	----	----	----	----	----
6-10	----	----	----	----	----
11-15	2	0.009	----	0.000	0.009
16-20	1	0.014	----	0.000	0.014
21-25	1	0.018	----	0.000	0.018
26-30	5	0.041	2	0.100	0.059
31-35	1	0.046	14	0.800	0.754
36-40	2	0.055	4	1.000	0.945
41-45	4	0.074	----	1.000	0.926
46-50	9	0.115	----	1.000	0.885
>50	192	1.000	----	1.000	0.000
Total	217		20		

Maximum difference = 0.945 95 percent critical K-S value = 0.318

time from vehicle violation to train arrival. There were a large number of violations especially at those locations that were not equipped with CWT systems. Inspection of Table 14 indicates that the majority of these violations occurred when the amount of warning time exceeded 50 sec. This occurred even at those locations where motorists had to drive around the gates. There is a definite increase in the number of violations for crossings with flashing lights and no CWT when the total warning time exceeds 35 sec.

A summary of the amount of time remaining from vehicle violation (the rear of the vehicle clearing the tracks) until the train entered the crossing is given in Table 15. It is interesting to note that five of these observations included clearance times of less than 6 sec.

The Kolmogorov-Smirnov two-sample tests was used to determine if the violations observed at crossings equipped with CWT systems exhibited the same population characteristics as those obtained at crossings without CWT systems. The analysis was performed by comparing crossings with similar types of warning devices. The analyses for violations occurring within categories of total warning time are given in Tables 16 and 17. Similar analyses for violations by time before train arrival are given in Tables 18 and 19. Each of these tests indicates that at the 95 percent level of confidence, there are significant differences between crossings with comparable types of warning devices, with and without CWT. CWT systems reduce the number of violations and, because they provide a more uniform amount of warning time, result in a greater proportion of violations occurring with smaller clearance time (interval of time between a vehicle clearing the tracks and the time of train arrival) than at crossings without CWT systems. The majority of vehicles that violate the warning devices at crossings

TABLE 17 KOLMOGOROV-SMIRNOV TEST ON THE NUMBER OF VIOLATIONS OCCURRING WITHIN CATEGORIES OF ADVANCE WARNING TIME (sec) FOR CROSSINGS EQUIPPED WITH FLASHING LIGHTS

Total Warning Time Interval	Flashing Lights With CWT		Flashing Lights Without CWT		Absolute Differences in Cumulative Occurrences
	Occurrences	Cumulative Occurrences	Occurrences	Cumulative Occurrences	
21-25	3	0.008	----	0.000	0.008
26-30	7	0.027	33	0.351	0.324
31-35	6	0.043	30	0.670	0.627
26-40	25	0.111	27	0.957	0.846
41-45	41	0.222	4	1.000	0.778
46-50	22	0.282	----	1.000	0.718
>50	265	1.000	----	1.000	0.000
Total	369		94		

Maximum difference = 0.846   95 percent critical K-S value = 0.157

TABLE 18 KOLMOGOROV-SMIRNOV TEST ON TIME (sec) FROM VEHICLE VIOLATION UNTIL TRAIN ARRIVAL FOR CROSSINGS EQUIPPED WITH GATES

Time from Violation Until Train Arrival	Gates Without CWT		Gates With CWT		Absolute Differences in Cumulative Occurrences
	Occurrences	Cumulative Occurrences	Occurrences	Cumulative Occurrences	
0-5	----	----	3	0.150	0.150
6-10	3	0.014	2	0.250	0.236
11-15	13	0.074	4	0.450	0.376
16-20	13	0.134	4	0.650	0.516
21-25	18	0.217	6	0.950	0.733
26-30	17	0.295	1	1.000	0.705
31-35	11	0.346	----	1.000	0.654
36-40	20	0.438	----	1.000	0.562
>40	122	1.000	----	1.000	0.000
Total	217		20		

Maximum difference = 0.733   95 percent critical K-S value = 0.318

TABLE 19 KOLMOGOROV-SMIRNOV TEST ON TIME (sec) FROM VEHICLE VIOLATION UNTIL TRAIN ARRIVAL FOR CROSSINGS EQUIPPED WITH FLASHING LIGHTS

Time from Violation Until Train Arrival	Flashing Lights Without CWT		Flashing Lights With CWT		Absolute Difference in Cumulative Occurrences
	Occurrences	Cumulative Occurrences	Occurrences	Cumulative Occurrences	
0-5	1	0.003	1	0.011	0.008
6-10	17	0.049	4	0.053	0.004
11-15	34	0.141	13	0.191	0.050
16-20	30	0.222	26	0.468	0.246
21-25	35	0.317	20	0.681	0.364
26-30	38	0.420	19	0.883	0.463
31-35	29	0.499	10	0.989	0.490
36-40	29	0.577	1	1.000	0.423
>40	156	1.000	----	1.000	0.000
Total	369		94		

Maximum difference = 0.490   95 percent critical K-S value = 0.157

equipped with CWT systems are, therefore, exposed to an increased probability of being struck by a train than violators at crossings without CWT systems. However, the number of violators is much smaller at crossings with CWT systems.

CONCLUSIONS

CWT systems are effective in providing a uniform warning time and in reducing motorist violations of the activated warning devices at the crossing.

The comparative analysis of vehicle-train accidents occurring from 1980 through 1984 indicated that crossings equipped with CWT systems have a lower accident rate than crossings without CWT. This difference was not, however, large enough to be statistically significant at the 95 percent confidence level.

REFERENCES

1. R. L. Monroe, D. K. Munsell, and J. T. Rudd. *Constant Warning Time Concept Development for Motorist Warning at Grade Crossings.* FRA, U.S. Department of Transportation, 1981.
2. *Rail-Highway Grade Crossing Accident/Incident and Inventory Bulletin, No. 6.* FRA, U.S. Department of Transportation, 1984.

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

Accident Causation Analysis at Railroad Crossings Protected by Gates

JOHN A. HALKIAS AND LAURENCE BLANCHARD

The purpose of this study was to identify probable causes of and factors responsible for accidents occurring at railroad crossings protected by gates. Two important goals of this study were to (a) compare the results obtained for the two types of warning systems activating the gates: fixed distance and constant warning time systems, and (b) test the hypothesis that extended, or widely variable, warning times create a lack of credibility in warning signals. These objectives were achieved by statistically analyzing accident data obtained from the National Rail-Highway Crossing Inventory and the Railroad Accident/Incident Report files for the period 1975 through 1984. An accident classification by circumstance (movement and position of the car in relation to the tracks and the trains) highlighted some causes and factors responsible for the different types of accidents. The classification indicated results and led to the development of a similar interpretation of the accidents for both types of warning systems. Further analysis confirmed and quantified the small impact of environmental factors (bad weather, poor visibility at crossings, etc.). Trends found in relation to warning times tended to indicate that lack of credibility in warning signals was a factor in the accidents.

More than 7,000 accidents involving grade crossings occur each year in the United States. They are responsible for approximately 600 fatalities and 2,500 injuries annually *(1)*. The high ratio of fatalities and injuries to the number of accidents at rail-highway grade crossings ranks these accidents among the most severe in the public safety area. As a reference, this ratio is approximately 40 times greater than the same ratio for all motorist accidents *(2)*.

In an attempt to reduce railroad crossing accidents, warning devices have been installed on or adjacent to the highway approaches to railroad grade crossings. These devices can be classified as either passive or active. Passive devices include stop signs, crossbucks, and pavement markings. They are used to direct attention to the location of the crossing and thus permit motorists to take appropriate action. Active devices include flashing lights and gates (automatic gates include flashing lights as a part of the warning display) that are train activated. They inform the motorist of the approach or presence of trains at grade crossings.

It should be pointed out that automatic gates are the most sophisticated and restrictive of all the grade crossing control devices: when activated, gates physically separate motor vehicles from the grade crossing. However, while 8 percent of public grade crossings are protected by gates, these crossings still account for about 15 percent of all train accidents involving grade crossings *(1)*. If this disproportion may be partly explained by high exposure (crossings having higher train and vehicle volume are usually equipped with automatic gates), it is still clear that a desirable safety level has yet to be achieved and more research is needed to investigate causes of these accidents.

J. A. Halkias, Department of Civil Engineering, Florida Institute of Technology, Melbourne, Fla. 32901. L. Blanchard, Matra Transport, Paris, France.

WARNING DEVICES

Two basic types of automatic control systems exist at crossings protected by gates: (a) fixed distance warning system, and (b) constant warning time system.

With a fixed distance warning system, trains activate the flashing lights and the gates at a predetermined distance from the crossing. This distance is calculated by using the speed of the fastest train allowed over the crossing and a specified minimum warning time. The major drawback of such systems is that the warning devices operate continuously while the train is on the approach track circuit, regardless of train speed. This leads to inconsistent warning time lengths for crossings used by trains having a wide speed range. Lengthy time intervals (e.g., slow train) between the signal activation and the arrival of the train at a crossing may lead to loss of credibility. Drivers may become impatient in situations where the warning device is active for a long time. Such repeated experiences can lead them to disregard the signals and to maneuver around crossing gates.

Signals activated by a constant warning time system do not present such a drawback. Constant warning time equipment has the capability of sensing a train in the approach section, measuring its speed and distance from the crossing, and activating the warning devices. Thus, regardless of train speed a uniform warning time is provided.

Many studies include lack of signal credibility as a factor in accidents and recommend equipping gates with a constant warning time system. Studies by Wilde et al. *(3)* and Halkias and Eck *(4)* provide useful information and recommendations for further analysis. Wilde et al. studied driver behavior at six crossings protected by gates activated by a fixed distance system *(3)*. Analysis of warning times at those crossings indicated an extreme variability from alarm period to alarm period as defined by standard deviations. The most variable warning times (ranging from 50 to 205 sec) occurred at a crossing at which several accidents involving train-vehicle collisions had occurred in the past. From this, the authors concluded that it can be speculated that the more variable the warning time, the higher the frequency of train-vehicle collisions.

By comparing accident rates at crossings before and after upgrading from a fixed distance to a constant warning time system, Halkias and Eck *(4)* found a 28 percent reduction in the accident rate. They concluded that this result tends to confirm the hypothesis that constant warning time systems have greater credibility than do fixed distance systems.

In-depth analysis of warning times appears necessary to test the hypothesis that extended or widely variable warning times contribute to accidents because they create a loss of credibility of the warning signals. Furthermore, although some studies *(3, 5)* have analyzed drivers' behavior at crossings protected by gates and deduced from it possible responsible factors, there are no statistical analyses on the circumstances and causes of the accidents. For this reason, and because it was believed that comparison of accident causes for the two types of warning systems (fixed distance warning and constant warning time systems), as well as any trend related to warning times would be more significant if studied on a large data sample, a statistical approach on a large data base was used for this study.

METHODOLOGY

The National Rail-Highway Crossing Inventory file and the Railroad Crossing Accident/Incident data file for the period January 1, 1975, through December 31, 1984, were obtained from the Federal Railroad Administration on six magnetic tapes. These two files were merged and correlated with the crossing identification number.

Two subfiles were extracted that contained all the accidents that occurred at railroad crossings protected by gates activated by both fixed distance warning system and constant warning time system. These subfiles were analyzed separately using the same procedure (except for the study concerning credibility factor, which was not applicable in the case of the constant warning time device).

Data Analysis

The analysis was divided into two parts: an accident classification by circumstance and an analysis of the accidents that remained unexplained by the circumstances.

Accident Classification by Circumstance

All the parameters available on the motorist action, the relative position of the car with respect to the train or trains (in a case of a second train), and the car movement when the impact occurred, were used to classify the accidents by circumstance. Table 1 gives the list of the parameters used and the type of information they provide. This classification also includes the interpretation of the accidents, the causes of which were understandable by the circumstances.

The data elements (or parameters) are those directly available from the inventory or accident/incident data files. Their definitions are given according to the procedures manual *(6)* and the FRA Guide for Preparing Accident/Incident Reports *(7)*.

TABLE 1 CIRCUMSTANCE PARAMETERS

Parameter	Values
Motorist	Drove around or through the gate Stopped and then proceeded Did not stop Other Unknown
Car position	Stalled on crossing Stopped on crossing Moving over crossing Unknown
Second train: The motorist drove behind or in front of a train and struck or was struck by a second train.	Yes No Unknown
Position of car / unit train: Identify the position within the train of the first locomotive unit or car that struck or was struck by the higher user.	

Analysis of the Accidents Unexplained by the Circumstances

Accidents in which the causes could not be explained by the circumstances were further analyzed to obtain additional information. These accidents were examined by using all of the available parameters that might highlight the causes of the accidents.

It should be pointed out that one important factor was missing—no human factor data were directly available from the file. Indeed, the only data elements available about the motorist were related to the action or more precisely to the movement of the car before impact. Information such as driver's age, location of residence (from which might be inferred the driver's familiarity with the crossing), and condition (whether or not intoxicated?) would have been helpful for a better understanding of the accidents. For this reason, only two parameters available were analyzed: (a) environmental conditions and (b) warning times.

Environmental Conditions

All the environmental elements that might have been contributors to the accident (wet roadway, poor visibility at the crossing, etc.) were regrouped in this category. For each element, a weight (w) of either 0, 0.5, or 1 was given to the probable adverse effect of the driver's response: 0 corresponding to a lack of adverse effect (good environmental conditions at the crossing) and 1 corresponding to a possible adverse effect. For each accident, a summation on (w) for all elements was calculated. An accident obtaining a total of 0 could be considered as not being adversely affected by any of these environmental conditions. The elements used and their corresponding weights are given in Table 2. A weight of 0.5 was given for fog or rain because these weather conditions were judged less dangerous than sleet or snow. Indeed, even if visibility is poor in case of fog or rain, stopping distance will not be severely affected as

TABLE 2 ENVIRONMENTAL PARAMETERS

Element	Weight
View of Track[a]	
Not obstructed	W = 0
Obstructed (permanent structure, standing railroad equipment, topography, highway vehicles, vegetation, other)	W = 1
Weather	
Clear or cloudy	W = 0
Fog or rain	W = 0.5
Sleet or snow	W = 1
Visibility	
Day	W = 0
Dawn, dusk: crossing illuminated by street lights or special lights:	
Yes	W = 0
No	W = 1
Dark: crossing illuminated by street lights or special lights:	
Yes	W = 0
No	W = 0.5

[a]Indicates if the driver's view approaching the crossing was obstructed to the extent that he/she might have been aware that a train was about to occupy or was occupying the crossing.

opposed to sleet or snow conditions where the motorist can completely lose control of the vehicle.

When the crossing was not illuminated by street lights or special lights, dark conditions received a weight of only 0.5 because they were considered less dangerous than dawn or dusk conditions. Red flashing lights offer more contrast with black background and are thus more conspicuous and visible. Furthermore, visibility at twilight is likely to be more diminished because of the continuously changing luminosity and the associated need for visual adaptation. When the crossing was illuminated by street lights or special lights, dark, dawn, and dusk conditions were judged similar to day conditions (weight of 0).

Warning Times

Warning times were obtained from the following data elements:

Maximum timetable speed (*MxTTSp*): The maximum train speed permitted over a crossing;

Typical minimum train speed (*MinSp*): The typical minimum train speed over a crossing; and

Train speed (*TrnSp*): The train speed when the accident occurred.

The concept of a fixed distance warning device is the provision of minimum warning time (*MinWT*) for the fastest train speed (*MxTTSp*) permitted over the crossing. To accomplish this requirement a train detection track circuit system is placed at a certain distance (*d*) from the crossing such that

$$d = MxTTSp \times MinWT \tag{1}$$

The minimum warning time (*MinWT*) corresponds to the interval of time between the arrival of the train at the track circuit—beginning of the signal's activation—and the arrival of the train at the crossing, for the case of a train traveling at the maximum timetable speed. This minimum warning time should be long enough to enable vehicles to stop or clear the crossing *(8)*. It was set to 24 sec for all of the crossings.

Furthermore, with maximum timetable speed and the typical minimum train speed known for each crossing, it is possible to work out the typical maximum warning time (*MaxWT*), which corresponds to the same interval of time as defined previously, but for a train traveling at the typical minimum speed.

$$d = MxTTSp \times MinWT$$

$$d = MinSp \times MaxWT \tag{2}$$

Hence

$$MaxWT = (MxTTSp)/(MinSp) \times MinWT \tag{3}$$

$$MaxWT = (MxTTSp)/(MinSp) \times 24 \text{ sec} \tag{4}$$

From the warning times a ratio was developed, the object of which was to enable the plotting of the actual warning time (*WT*) with respect to the minimum and maximum warning times (*MinWT* and *MaxWT*) for each accident. The following scheme explains the calculation of the ratio. Considering a line with three points A, B, and X of respective coordinates *MinWT*, *MaxWT*, and *WT*;

____*A*__________*X*___________________*B*____

the location of *X* is given by the parameter *r* such that

$$AX = r \times AB \tag{5}$$

$$r = AX/AB \tag{6}$$

Because

$$AX = WT - MinWT \tag{7}$$

$$AB = MaxWT - MinWT \tag{8}$$

$$r = (WT - MinWT)/(MaxWT - MinWT) \tag{9}$$

As an example, $r = 0$ corresponds to an actual warning time equal to the minimum warning time, and $r = 1$ corresponds to an actual warning time equal to the maximum warning time. The sketch below shows the relative scale adopted and the plotting of two actual warning times (*WT*1 and *WT*2 as examples):

0 ____·________·________·____________·______

MinWT	*WT*1	*MaxWT*	*WT*2	
$r = 0$	$r = 1/2$	$r = 1$	$r = 2$	(10)

The idea was to highlight the credibility problem, if any, by finding a correlation or trend between the frequency of accidents and the warning times calculated or the ratio developed, or both.

It should be noted that this procedure is not applicable in the case of constant warning time systems because, regardless of the train speed, the warning time remains constant and, in general, equals 24 sec.

RESULTS

The results obtained for both types of warning devices (fixed distance and constant warning time devices) were similar and led to identical classifications and interpretations of the accidents (except for the accidents related to the credibility factor). The data in Tables 3 and 4 show the distribution of accidents for both types of warning devices. The data in Table 5 and Figure 1 compare the results obtained for the two types of warning devices by using the valid percentage.

TABLE 3 ACCIDENT CLASSIFICATION BY CIRCUMSTANCE-FIXED DISTANCE WARNING DEVICES

Circumstances	Frequency	Percent
Accidents remaining unexplained	2,585	43
Stopped	1,058	18
Stalled	1,136	19
Struck or was struck by a second train	175	3
Struck a car other than the leading car	451	8
Missing or unknown	569	9
Total	5,974	100

TABLE 4 ACCIDENT CLASSIFICATION BY CIRCUMSTANCE (CONSTANT WARNING TIME DEVICES)

Circumstances	Frequency	Percent
Accidents remaining unexplained	723	36
Stopped	342	17
Stalled	359	18
Struck or was struck by a second train	54	3
Struck a car other than the leading car	135	7
Missing or unknown	375	19
Total	1,988	100

- Twenty percent of the accidents that occurred at railroad crossings with fixed distance warning systems (21 percent for constant warning time systems) involved motorists stopped on the crossing. The presence of a highway intersection within 75 ft of the crossing was found to contribute to these accidents. Another factor believed responsible for these accidents was motorists' lack of driving experience.
- Twenty-one percent of the accidents that occurred at railroad crossings with fixed distance warning systems (22 percent for constant warning time systems) involved motorists stalled on the crossing. The cause of these accidents was believed to be engine failure or automobile malfunction.
- Eight percent of the accidents that occurred at railroad crossings with fixed distance warning systems (9 percent for constant warning time systems) involved motorists who drove around or through the gate and struck a train car other than the leading car. The environmental factor did not have a strong adverse effect because more than 50 percent of these accidents occurred during good weather and good visibility conditions (Tables 6 and 7). Alcohol and drug intoxication, brake failure, or inattentiveness were believed to contribute to these accidents.
- Forty-eight percent of the accidents that occurred at railroad crossings with fixed distance warning systems (45 percent for constant warning time systems) involved motorists who drove around or through the gate, and while moving over the crossing, struck or were struck by the leading train car (accidents classified as remaining unexplained by the circumstances). This last category of accidents was further analyzed

TABLE 5 COMPARISON OF ACCIDENT CLASSIFICATION BY CIRCUMSTANCE (VALID PERCENT) FOR FIXED DISTANCE AND CONSTANT WARNING TIME DEVICES

	Valid Percent	
Circumstance	Fixed Distance	Constant Warning Time
Accidents remaining unexplained	48	45
Stopped	20	21
Stalled	21	22
Struck or was struck by a second train	3	3
Struck a car other than the leading car	8	9

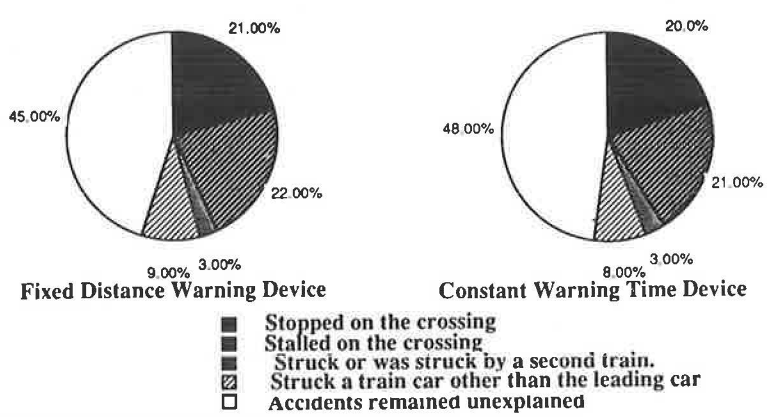


FIGURE 1 Comparison of the accident classification by circumstance for fixed distance and constant warning time systems.

TABLE 6 ACCIDENTS BY ENVIRONMENTAL FACTOR WEIGHT (VEHICLE STRUCK A TRAIN CAR OTHER THAN THE LEADING CAR, FIXED DISTANCE WARNING DEVICES)

Weight	Frequency	Percent	Valid Percer
0.0	168	37	51
0.5	85	19	26
1.0	58	13	18
1.5	12	3	4
2.0	3	1	1
Missing	125	27	Missing
Total	451	100	100

TABLE 7 ACCIDENTS BY ENVIRONMENTAL FACTOR WEIGHT (VEHICLE STRUCK A TRAIN CAR OTHER THAN THE LEADING CAR, CONSTANT WARNING TIME DEVICES)

Weight	Frequency	Percent	Valid Percent
0.0	61	45	54
0.5	22	16	19
1.0	19	14	17
1.5	8	6	7
2.0	1	1	1
2.5	3	2	2
Missing	21	16	Missing
Total	135	100	100

TABLE 8 ACCIDENTS BY ENVIRONMENTAL FACTOR WEIGHT (FIXED DISTANCE WARNING DEVICES)

Weight	Frequency	Percent	Valid Percent
0.0	826	32	47
0.5	363	14	21
1.0	445	17	26
1.5	63	2	4
2.0	39	2	2
2.5	4	0	0
Missing	845	33	Missing
Total	2,585	100	100

TABLE 9 ACCIDENTS BY ENVIRONMENTAL FACTOR WEIGHT (CONSTANT WARNING TIME DEVICES)

Weight	Frequency	Percent	Valid Percent
0.0	295	41	51
0.5	104	14	18
1.0	150	21	26
1.5	16	2	3
2.0	10	1	2
Missing	148	21	Missing
Total	723	100	100

by examining the environmental conditions and the warning times.

Environmental Conditions

For approximately 30 percent of these accidents (results similar for both types of warning systems), bad weather or poor visibility at crossings might have had a likely or strong adverse effect (weight larger or equal to 1) on motorists' action and decision (Tables 8 and 9). Although the relative weights given to some environmental factors such as fog and rain versus sleet and snow and variable luminosity can be questioned, the results indicate that a small change in the relative weights would not have affected the conclusion that most of the accidents occur during good weather and good visibility conditions.

Warning Times

The >>*x*% sign to the right of Figures 2 and 3 indicates that *x* percent of the accidents had an actual warning time larger than the extreme value plotted. These cases were largely spread on the time scale, and for reasons of scale, do not appear on the figures.

For fixed distance, several conclusions were drawn from the analysis of warning times.

- Warning times have an extreme variability. They range from less than 20 sec up to 16 min (Figure 2). This variability is much larger than the one found by Wilde et al. *(3)*. It should be noted that their research was based on the study of only six crossings. It can be assumed that these few crossings were not representative of the whole crossing population.
- Three percent of the accidents occurred because of warning times that were too short (smaller than the minimum warning time) to enable motorists to clear the crossing before the arrival of the train (Figures 2 and 3). This was believed to be a result of the introduction of high-speed rail service at existing facilities without any corrective action having been undertaken to provide a minimum required warning time at the crossing. It should be noted that this percentage might have been biased by incorrect data, such as maximum timetable speed smaller than the true value or overestimation of the actual train speed.
- Comparing different groups of crossings (classification based on the value of the typical maximum warning time), a general trend was found: the larger the typical maximum warning time, the less the accidents are spread on the time scale, and the more concentrated they are near the minimum warning time. In other words, the more variable the warning times (the larger is the train speed range), the more accidents occur when the actual warning time is short and close to the minimum warning time (e.g., actual train speed close to the maximum timetable speed). Also, the smaller the warning time range, the more accidents occur when the actual warning time is large and beyond the typical maximum warning time (actual train speed smaller than the typical minimum speed).

The distributions obtained for the first two charts in Figure 4 (maximum warning times less than 0.75 min and maximum warning time between 0.75 and 1.5 min) are unexpected in terms of the large percentage of accidents occurring out of the typical warning time range (61 and 48 percent, respectively). Indeed, because most of the trains can be expected to provide a

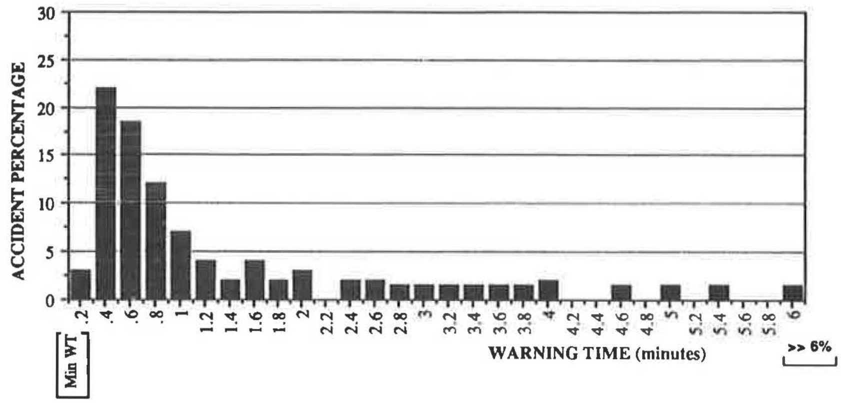


FIGURE 2 Accident percentage versus actual warning time.

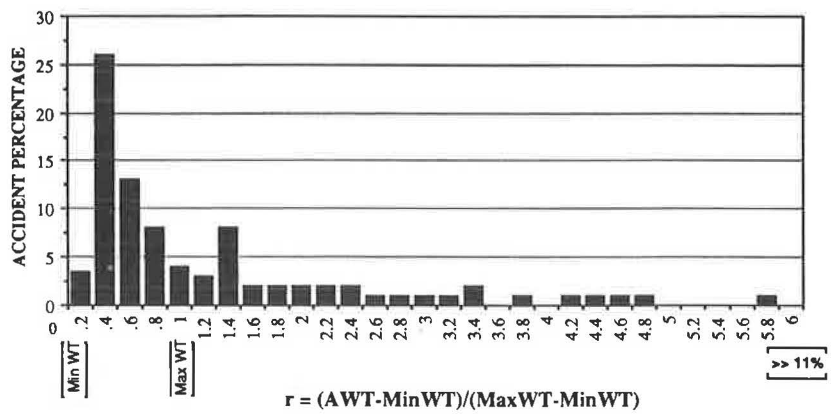


FIGURE 3 Accident percentage distribution with respect to the minimum and typical maximum warning time.

warning time between the minimum and the typical maximum times, most of the accidents should be expected to occur within this range of warning times. From this result, it can be inferred that, in the case of crossings providing a small range of warning times (e.g., typical minimum speed close to the maximum timetable speed), larger warning times are dangerous. These large warning times, being out of the typical range, are by definition infrequent. They are, however, involved in a high percentage of accidents. If it is considered that the driver involved in the accident was familiar with the crossing *(2, 9, 10)*, it is likely that he experienced short warning times at the crossing. A longer alarm period without an approaching train might have led drivers to distrust the signal, and, getting impatient, they might have proceeded without looking for a train.

The trend leading to a concentration of accidents when the typical maximum warning time increases is optimal for the last chart in Figure 4 (maximum warning time greater than 6 min), with 40 percent of accidents occurring when the actual warning time is close to the minimum warning time. Referring to the probable familiarity of the driver with the crossing, this trend can be explained by a lack of driver trust in the signal. The larger the typical maximum warning time, the higher the probability that the driver familiar with the crossing has experienced a long alarm period at the crossing. The driver might have finally developed the expectancy to have to wait a long time at the crossing and decided that there was enough time to proceed before the arrival of the train. Changes in the train pattern (faster train providing a shorter warning time) may be responsible for this large number of accidents.

Although it was impossible to quantify the importance of the lack of credibility in warning signals, the trends found indicate that it is a factor in accidents.

In the case of accidents occurring at gates activated by a constant warning time system, the credibility factor should not be involved because the warning time provided at these crossings is short, constant, and approximately 24 sec. However, crossings equipped with fixed distance warning systems are much more numerous than crossings equipped with constant warning time systems. In 1984, 12,483 crossings were equipped with fixed distance warning systems compared with 3,953 crossings equipped with constant warning time systems. The probability that a driver will encounter a crossing equipped with a fixed distance warning system is therefore much higher. Furthermore, drivers do not have any knowledge about warning devices that would enable them to differentiate between crossings equipped with constant warning time systems and those equipped with fixed distance warning systems. Thus drivers might in some cases have carried over their expectancy of extended or inconsistent warning times developed at cross-

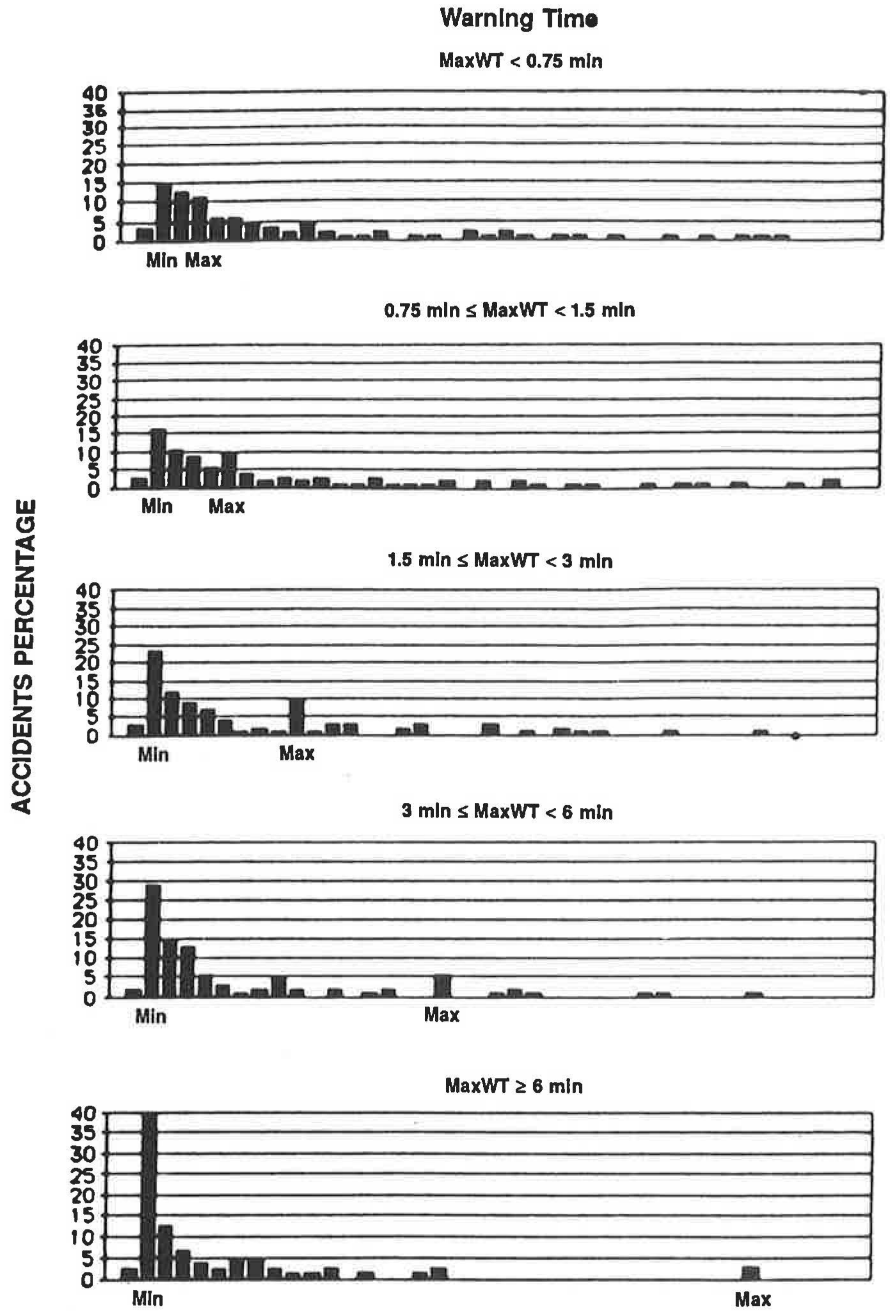


FIGURE 4 Summary of accident percentage distribution for five ranges of typical maximum warning times.

ings equipped with fixed distance warning systems to crossings equipped with constant warning time systems. In other words, it might be inferred that the effectiveness of constant warning devices is influenced by the inconsistency of warning times at the numerous crossings equipped with fixed distance warning systems. In this case, the warning signal's lack of credibility would also be a factor in accidents occurring at crossings equipped with constant warning systems. The similarity of the results for both types of warning systems and the high percentage of accidents remaining unexplained by the circumstances and environmental factors, would tend to confirm this theory. An analysis of driver familiarity with crossings equipped with constant warning time systems might have allowed this problem to be highlighted; for drivers involved in these accidents, a lack of familiarity with such crossings will reinforce the hypothesis developed.

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

This study highlighted some causes of and factors that are responsible for accidents at railroad crossings. Unlike previous studies, it analyzed a large data base and provided primary statistical data on accidents occurring at railroad crossings protected by gates. The study might be useful as a base for the development of necessary countermeasures for improving safety at railroad crossings protected by gates.

Some of the conclusions reached are as follows:

- Results from the present large data sample confirm the theory generally adopted that the majority of accidents occurs during good weather and good visibility conditions.
- Physical and environmental conditions are not sufficient to explain accidents. For a large percentage of cases, the cause of the accidents remained obscure or uncertain essentially because of a total lack of data elements on the driver.
- Study of warning times led to two main conclusions. (a) Inconsistency in warning time length leads motorists to distrust signals. At railroad crossings that have a narrow typical warning time distribution, most of the accidents occur beyond the typical maximum warning time. (b) Extended warning times lead motorists to ignore warning signals and cross the railroad.

From these results, it was concluded that lack of credibility in warning signals was a factor in accidents occurring at crossings equipped with fixed distance warning systems.

For crossings equipped with constant warning time systems, it was hypothesized that the effectiveness of this warning device was biased by the inconsistency of warning times at the numerous crossings equipped with fixed distance warning systems. From this, it was concluded that the warning signals' lack of credibility might also contribute to the accidents occurring at crossings equipped with constant warning time systems.

Recommendations

Based on the work undertaken for this research, the study of the data sources, and the results obtained, several suggestions are presented. They concern two important subjects: (a) the data available in the U.S. Department of Transportation Crossing Inventory and the FRA Rail-Highway Crossing Accident/Incident files, and (b) areas in which further research could prove helpful.

The following recommendations are made:

- For better quality and reliability in the data, it is important to minimize inconsistencies. This can be achieved by running programs to check the consistency of the values entered in the data file. Programs as simple as the one that checks whether a field declared as numerical contains only digits might enable the correction of mistyped characters. Other checking programs should detect incompatibilities such as a vehicle stopped on the crossing with a speed other than zero.
- More precise information is needed on the motorist's action when the accident occurred. The phrase "motorist drove around or through the gate" has to be reviewed. For a better understanding of the accidents, it is important to be able to distinguish the cases in which the motorist drove around the gate from those in which he drove through the gate. Furthermore, these two motorist actions involve two different approaches of solving the problem. For example, a countermeasure to stop drivers from driving around the gates would be to install four half-gates (instead of two) to completely separate the motorist from the tracks. However, this countermeasure would probably not have any impact on the accidents in which the motorist drove through the gates. For these cases increasing the conspicuousness of the gates might improve safety.
- Information about the motorist is indispensable for accident causation analysis. Motorist data elements that may prove valuable are age, alcohol and drug intoxication, and location of residence (from which can be inferred driver familiarity with the crossing).
- Additional data on the train speed pattern at crossings (such as median train speed) will enable a more accurate definition of the warning time distribution at crossings and thus will provide more information on the effect of warning times on accidents.

Further research could prove valuable in two areas:

- A further causation analysis of these accidents is needed. This analysis should concentrate on the possible contributing factors about which information was not available in the data source used for this research. The important factors to be examined are alcohol and drug intoxication, advanced age, lack of driving experience, and automobile malfunction.
- The results obtained lead to the development of the hypothesis that the credibility factor might also be involved in accidents occurring at crossings protected by constant warning time systems. An analysis of driver familiarity at these crossings might provide valuable information. For drivers involved in the accidents at these crossings, a lack of familiarity with the crossings would reinforce the hypothesis developed.

REFERENCES

1. *Rail-Highway Crossing Accident/Incident and Inventory Bulletin 5.* FRA, U.S. Department of Transportation, 1983.
2. J. H. Sanders, G. S. Kolsrud, W. G. Berger. *Human Factors Countermeasures to Improve Highway-Railway Intersection Safety.* DOT-HS-190-2-300. U.S. Department of Transportation, 1973.
3. G. J. S. Wilde, L. J. Cake, and M. B. McCarthy. *An Observational Study of Driver Behavior at Signalized Railroad Crossings.* Report 75-16. Canadian Institute of Guided Ground Transport, Kingston, Ontario, 1975.
4. J. A. Halkias and R. W. Eck. Effectiveness of Constant Warning Time Versus Fixed Distance Warning Systems at Rail-Highway Grade Crossing. In *Transportation Research Record 1010*, TRB, National Research Council, Washington, D.C., 1985, pp. 101–116.
5. T. Ikeda. A View of Level Crossing Accident in JNR. *Japanese Railway Engineering*, No. 12, 1971, pp. 12–15.
6. *Procedures Manual, National Railroad-Highway Crossing Inventory.* FRA, U.S. Department of Transportation, 1973.
7. *FRA Guide for Preparing Accident/Incident Reports.* FRA, U.S. Department of Transportation, 1982.
8. *Railroad-Highway Grade Crossing Handbook.* FHWA, U.S. Department of Transportation, 1978.
9. D. W. Shoppert and D. W. Hoyt. *NCHRP Report 50: Factors Influencing Safety at Highway-Railway Grade Crossings.* TRB, National Research Council, Washington, D.C., 1968.
10. E. C. Wigglesworth. Epidemology of Road-Rail Crossing Accidents. *Journal of Safety Research*, Vol. 22, No. 4, 1979, pp. 162–171.

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

How to Estimate the Safety of Rail-Highway Grade Crossings and the Safety Effects of Warning Devices

E. HAUER AND B. N. PERSAUD

The safety of a rail-highway grade crossing should be estimated by mixing information about causal factors such as train and traffic flows, type of warning device, and geometry with the accident history of the site. A theoretically sound and logically coherent procedure for doing so is suggested. This procedure is applicable not only to crossings but also to intersections, road segments, drivers, and vehicles. Current estimates of the safety effect of warning devices used at crossings are incorrect. This claim is supported in this paper. The aforementioned coherent procedure for estimating safety also enables corrected estimates of effectiveness to be furnished, which should be used in the allocation of resources to warning devices.

Addressed in this paper are two fundamental questions that arise in the context of managing safety at rail-highway grade crossings: first, how to identify crossings that are "unsafe"; second, how to estimate the safety effect of warning devices. To answer the first question it is necessary to be in a position to estimate the safety of a crossing. To answer the second question it is necessary to estimate what the safety of an upgraded crossing would have been had the warning device not been changed and what the safety of the crossing is with the new warning device in place. For both questions the elemental task is the task of safety estimation. The same elemental tasks are required when the entity under consideration is not a crossing but an intersection of roads, a road section, a driver, or a vehicle. This is why matters of method discussed in this paper are of general interest.

To forestall the possibility of miscommunication it is best to define the term safety. Safety is considered a property of a specific entity; in this case, it is the property of a certain crossing. The safety property of a crossing is defined as the number of accidents and their adverse consequences expected to occur on a crossing per unit of time. The term expected means "what would be the average in the long run were it possible for all relevant conditions to remain unchanged." It should be noted that the term "safety" is certainly not equivalent to the count of accidents recorded on a crossing. The count of accidents is a reflection of the safety of the crossing and serves as a clue for the estimation of safety.

Safety Studies Group, Department of Civil Engineering, University of Toronto, Toronto, Ontario, Canada M561A4.

HOW TO ESTIMATE SAFETY

When asked to give an estimate of the safety of a specific crossing, the natural inclination is to ascertain the key features of the crossing (type of warning device used, prevailing train and traffic flows, character of the surrounding area, etc.) and base the answer on what is known to be the typical "accident experience" of crossings with similar characteristics. The accident experience of the crossing about which the inquiry is made does not appear to ordinarily affect the estimate. Of the 12 "common rail-highway crossing hazard models" listed by Farr *(1)*, 8 take this approach.

In contrast, were a traffic engineer asked about the safety of a certain intersection, he would likely base his answer on the history of accidents for that specific intersection. If a few years of accident history are available, it is not common to base the safety estimate on information about traffic flows and other causal factors.

Thus, in principle, two sources of information bear on the task of safety estimation: causal factors, which tell something about the safety of similar entities (crossings, intersections, etc.), and accident records, which capture the history of the specific entity, the safety of which is examined. It makes common sense, that neither source should be disregarded. Therefore, at least in principle, both examples of practice are lopsided and less than fully efficient in estimating safety.

Of course, the difference between crossings and intersections is that crossing accidents are very rare whereas intersection accidents occur relatively more often. Thus, in practical terms, perhaps not much is lost by disregarding in estimation the accident history of a crossing or the characteristics of the intersection. Still, to be sure, solid analysis is required.

Some crossing hazard models (e.g., the Oregon, Utah, Detroit, and Wisconsin formulae) do incorporate the accident experience of a crossing, usually as an additive term. For these, it is unclear why the accident record term should be simply added to the other variables and if so what should be the relative weights of the two diverse pieces of evidence. The DOT "General Accident Prediction Formula" *(2)* is, to the authors' knowledge, the first model to explicitly advocate that an estimate of crossing safety based on causal factors be linearly combined with an estimate of safety based on the accident history of the crossing. The approach is credited to Mengert *(3)* and draws on the techniques of empirical Bayesian analysis as does the discussion in the next section. Thus, the problem to be analyzed in the next section is how to combine data about traffic, geometry, and other causal factors with the accident history of a crossing so as to obtain an efficient and logically sound estimate of the hazard of the site.

THEORY AND EVIDENCE

The aim is to obtain an estimate of the number of accidents (perhaps by type and severity) expected to occur at a specific crossing during a certain period of time. This expected number will be denoted by m, and it will serve to measure the safety of the specific crossing under scrutiny. The safety of this crossing, m, will never be known precisely, and an estimate of m to be denoted by $\hat{m}$ will have to suffice. The estimate $\hat{m}$, which is sought here, is to be based on the causal factors that describe the crossing as well as on its recorded accident history. It is in this dual foundation that the approach to estimation advocated here differs from much of current practice.

Empirical evidence will be used to both motivate the theory and to support its validity. However, the main strength of the theory offered here is not that it fits accident data but that it is derived from plausible postulates by deductive reasoning. For this reason, it may have rather general validity and wide applicability.

The point of departure is the assertion that every crossing is characterized during a specific period of time by its own m. Even if a group of crossings are considered that are similar in terms of traffic, warning device used, and other measured characteristics, their m's will not be identical. There is always much about a crossing that is unique and remains unmeasured. For such a group of similar crossings, the mean of their m's will be denoted by $E\{m\}$ and the dispersion of their m's will be measured by their variance, to be denoted $VAR\{m\}$.

Let x denote the count of accidents for a specific crossing of this group during a certain period of time. Furthermore, let $n(x)$ denote the number of crossings in the group on which the accident count was x. It is not surprising that from the $n(0)$, $n(1)$, $n(2)$ . . ., which describe the count of accidents in this homogeneous group of crossings, something can be learned about both $E\{m\}$ and $VAR\{m\}$. It can be shown that if accident occurrence on each crossing obeys the Poisson probability law, then

$$E\{x\} = E\{m\} \tag{1a}$$

$$VAR\{x\} = VAR\{m\} + E\{m\} \tag{1b}$$

Equations 1a and 2b allow the assertion that the m's vary from crossing to crossing to be substantiated and $VAR\{m\}$ to be estimated. This will be done with reference to real data.

The data consist of a full description of some 200,000 crossings in the United States—their geometry, train and vehicle volumes, warning devices used, number of tracks, type of area, and, of course, complete record of accident occurrence (during the years 1980 to 1984). From this data set a fairly homogeneous group of crossings have been extracted (for purposes of illustration) with the following characteristics: urban area, 1 track, 1 to 2 trains per day, 0 to 1,000 vehicles per day, and equipped with crossbucks. Table 1 gives the number of such sites with 0, 1, 2 . . . accidents in each of the years 1980 to 1984.

From the data in Table 1, it is easy to calculate the sample mean $\bar{x}$ and sample variance s^2 of the number of accidents in a homogeneous group of crossings using the following equations:

$$\bar{x} = \sum x \cdot n(x)/\sum n(x) \tag{2}$$

$$s^2 = \sum[(x - \bar{x})^2 \cdot n(x)]/\sum n(x) \tag{3}$$

These two sample statistics can be used to estimate $E\{x\}$ and $VAR\{x\}$ in Equation 1. Thus, the standard deviation of the m's in a group of crossings can be estimated by

$$\hat{\sigma}\{m\} = (\hat{VAR}\{m\})^{1/2} = (s^2 - \bar{x})^{1/2} \tag{4}$$

Using Equations 2–4 on the data in Table 1, the entries of Table 2 are calculated.

Inspection of Table 2 reveals that the standard deviation of the m's in this group of crossings is perhaps twice the mean of the m's. This confirms the assertion that even within a fairly homogeneous group of crossings it may not be assumed that all m's are the same. Perhaps this conclusion should have been self-evident even without empirical substantiation.

It could be argued that the group of crossings used in this illustration is still not entirely homogeneous; that, were the group to consist of crossings that all carried, for example, 500

TABLE 1 ACCIDENTS OCCURRING IN A HOMOGENOUS GROUP OF CROSSINGS

Number of Accidents [x]	Number of Crossings with x Accidents [n(x)]				
	1980	1981	1982	1983	1984
0	9770	9796	9794	9790	9838
1	160	138	141	143	96
2	8	5	4	5	5
3	1	0	0	1	0

NOTE: Urban; 1 track; 1–2 trains/day; 0–1000 vehicles/day; crossbucks.

TABLE 2 SAMPLE MEAN ($\bar{x}$), SAMPLE VARIANCE (s^2), $\hat{VAR}\{m\}$ AND $\hat{\sigma}\{m\}$

	1980	1981	1982	1983	1984
$\bar{x}$	.0180	.0149	.0150	.0157	.0107
s^2	.0199	.0157	.0156	.0171	.0110
$\hat{VAR}\{m\}$	.0019	.0008	.0006	.0014	.0003
$\hat{\sigma}\{m\}$	.04	.03	.02	.04	.02

vehicles per day, the variance of m would approach 0. Computation based on analysis of variance relationships as well as empirical investigation of more homogeneous groupings show that this contention is not valid; that there is a limit beyond which the $VAR\{m\}$ cannot be reduced. At this point $VAR\{m\}$ reflects the differences between the entities of a group that are not captured by their measured characteristics. The point of departure was the assertion that in a group of crossings that are homogeneous in their measured characteristics, the m's have a distribution with a positive variance. This assertion is now assumed to have been substantiated.

At this point an assumption needs to be introduced that cannot be directly substantiated; only evidence can be shown that it allows good predictions. It is well known that if the distribution of m's within a group of crossings can be described by a Gamma probability distribution, the counts $n(x)$ should obey the negative binomial distribution. Listed in Table 3 are the observed counts $n(x)$ and what is predicted by the negative binomial distribution. The correspondence between the observed and the predicted values is embarrassingly close. For the present purpose it can be taken to mean that the assumption that the distribution of m's can be represented by the Gamma distribution is not contradicted by data.

It has been shown in a recent paper *(4)* that when on each crossing, accident occurrence is governed by the Poisson probability law, and when the distribution of m's in a group of crossings can be described by a Gamma probability distribution function, it is best to estimate the m of a crossing on which the count of accidents was x by the estimator ε:

$$\varepsilon = x + [E\{m\}/(VAR\{m\} + E\{m\})]\,[E\{m\} - x] \quad (5a)$$

$$= \alpha E\{m\} + (1 - \alpha)x \quad (5b)$$

where

$$\alpha = (1 + VAR\{m\}/E\{m\})^{-1}$$

In Equation 5, ε is the estimator of m for a crossing, which recorded x accidents and, which by its measured causal factors, belongs to a population of crossings in which the m's have a mean $E\{m\}$ and variance $VAR\{m\}$. This is the central result of this section. It shows that the estimator ε is always a mixture of what is observed (x) and of what is known to be the mean of the m's for the group to which the site belongs ($E\{m\}$). The "weight" (α) of $E\{m\}$ is always a number between 0 and 1. When the $VAR\{m\} >> E\{m\}$ (the group of crossings is very diverse in m's), α will be very small and $\varepsilon \cong x$. That is, little can be learned from the fact that a crossing has certain measured characteristics because crossings with similar characteristics differ widely in their m's. Conversely, when $VAR\{m\} << E\{m\}$ (the group of crossings is homogeneous in their m's), the weight $1 - \alpha$ will be very small and $\varepsilon \cong E\{m\}$. In this case, little weight attaches to x, which is given to random fluctuations, and

TABLE 3 OBSERVED AND PREDICTED COUNTS OF CROSSINGS

Number of	Number of Crossings* [n(x)] with [x] Accidents									
Accidents	1980		1981		1982		1983		1984	
[x]	Obs.	Pred.	Obs.	Pred.	Obs.	Pred.	Obs.	Pred.	Obs.	Pred.
0	9770	9770	9796	9796	9794	9794	9790	9791	9838	9838
1	160	159	138	139	141	141	143	141	96	97
2	8	9	5	4	4	4	5	7	5	4
3	1	1	0	0	0	0	1	0	0	0

NOTE: Urban; 1 track; 1–2 trains/day; 0–1000 vehicles/day; crossbucks.

one should rely mainly on the fact that crossings in this group all have similar m's.

Note that in Equation 5 the period over which x accidents have been counted must be of the same length as that to which $E\{m\}$ and $VAR\{m\}$ pertain. Consider a case in which what is known about the m's is for a period i years long whereas the count x has been obtained over a period of j years. If the m's do not change in time, $E_j\{m\} = (j/i)E_i\{m\}$ and $VAR_j\{m\} = (j/i)^2 VAR_i\{m\}$. The subscripts are used to designate the length of the period to which the m's pertain. Thus, for example, if $i = 2$, m has the dimension of accidents in 2 years. Now the weight to use in Equation 5 is given by

$$\alpha_j = [1 + (j/i)VAR_i\{m\}/E_i\{m\}]^{-1} \quad (6a)$$

and therefore

$$\varepsilon_j = \alpha_j(j/i)E\{m_i\} + (1 - \alpha_j)x_j \quad (6b)$$

It is instructive to note in Equation 6a that as the length (j) of the recorded accident history increases so α_j diminishes. Thus, the more that is known about the accident history of a crossing, the lesser is the weight attached to what happens at similar crossings, and the more the actual accident record is relied on for estimation. When the length of the accident history increases without limit, $\hat{m} = \bar{x}$. Conversely, when the recorded accident history is meager, most of the weight attaches to what is gleaned from the accident record of similar sites. In the limit, when no accident history exists, the best estimate is the mean value for similar sites.

It remains to be seen what should be used for $E\{m\}$ and $VAR\{m\}/E\{m\}$ in Equations 5 and 6. Two distinct options are considered. One option is to create groups of crossings that are similar and to estimate $E_i\{m\}$ by $\bar{x}_i$ and $VAR_i\{m\}$ by $s_i^2 - \bar{x}_i$. This makes Equations 5 and 6 into

$$\hat{m}_j = \hat{\alpha}_j(j/i)\bar{x}_i + (1 - \hat{\alpha}_j)x_j \quad (7)$$

$$\hat{\alpha}_j = [1 + (j/i)\,(s_i^2 - \bar{x}_i)/\bar{x}_i]^{-1} \quad (8)$$

To illustrate, consider the following example. A grade crossing equipped with crossbucks is in an urban area, has one track, two trains per day, carries 800 vehicles per day, and in the last 10 years has recorded two accidents. The safety of that crossing can be estimated as follows. From annual accident counts on a large group of similar crossings the sample mean is 0.015 accidents per year and the sample variance is 0.017 (accidents per year)^2. Thus, $i = 1$ year while $j = 10$ years and $\hat{\alpha}_{10} = [1 + 10 \times 0.002/0.015]^{-1} = 0.43$. It follows that the estimate of m_{10}, $\hat{m}_{10} = 0.43 \times (10 \times 0.015) + 0.57 \times 2 = 1.2$ accidents in 10 years or an annual rate of 0.12 accidents.

Ten years of accident data are used in this example for purposes of illustration; no suggestion is implied about what should be proper practice. The natural inclination is to ask whether in view of the many changes that occur during a decade, the use of a shorter accident history may not be appropriate. It can be countered that at a crossing at which an accident occurs on the average once in 67 years (1/0.015), 10 years of accident history is too short. This question needs to be explored further.

The second distinct option for estimating $E\{m\}$ and $VAR\{m\}/E\{m\}$ is to make use of the results of multivariate analysis. After all, $E\{m\}$ is the "mean of the m's" for crossings with specified characteristics. This is precisely what multivariate models are supposed to estimate. Similarly, the "residuals" that are a by-product of multivariate modeling contain information about the $VAR\{m\}$.

A "generalized linear modeling" software package GLIM *(5)* has been used for this purpose. This appears to be the preferred approach for several reasons. First, it allows the representation of accident counts as coming from a negative binomial distribution. Second, it yields maximum likelihood estimates of the parameters.

The data consisted of the 1980 to 1984 accident records of some 200,000 crossings in the United States and information about train traffic, vehicular traffic, and other crossing characteristics that have been extracted from the grade crossing inventory file. Crossings were classified into eight groups by warning device (crossbucks or flashers), type of setting (urban or rural), and number of tracks (single or multiple). The same classification has been used earlier by Coleman and Stewart *(6)*. For each group of crossings, parameters of a model equation were estimated. The model equation is of the form

$$\ln[E\{m\}] = b_0 + b_1[\ln(C)] + b_2[\ln(T)] + b_3[\ln(T)]^2 \quad (9)$$

In this equation, $E\{m\}$ is for the 5-year period 1980 to 1984; the four parameters b are estimated by GLIM; C stands for average annual daily traffic, and T is the total number of through trains per day.

This model form has been chosen after several trials and is consistent with the Coleman and Stewart models *(6)* and the U.S. Department of Transportation "best volume" accident prediction formula *(3)*. The intent was to obtain a proper basis for the estimation of $VAR\{m\}$ and not to improve on the DOT basic formula, which contains many more parameters. Table 4 gives the parameter estimates for Equation 9.

To estimate $VAR\{m\}$, use has been made of the fact that (if the model equation is correct) each squared residual can be regarded as an estimate of $VAR\{x\}$. Figure 1 shows a plot of means of squared residuals for groups of crossings that have similar estimates of $E\{m\}$. Such plots as well as the literature *(7)* suggest a relationship of the form:

$$VAR\{x\} = E\{m\} + [E\{m\}]^2/k \quad (10)$$

In view of the relationship in Equation 1b, Equation 10 implies that $VAR\{m\}$ is proportional to the square of $E\{m\}$ with $1/k$ as the coefficient of proportionality. (This notation is consistent with what is used in GLIM.) Maximum likelihood estimates of k for Equation 10 are given in the following table:

	Crossbucks		*Flashers*	
	Rural	*Urban*	*Rural*	*Urban*
Single track	0.48	0.52	0.66	0.74
Multiple track	0.62	0.54	0.49	0.65

To illustrate the use of these results, consider again the crossing equipped with crossbucks: it is in an urban setting, has one track, two trains per day, carries an average of 800 vehicles

TABLE 4 PARAMETERS FOR MODEL EQUATION

			b_0	b_1	b_2	b_3
Crossbucks,	Rural,	Single Track	-6.078	0.524	0.966	-0.092
"	"	Multiple Tracks	-5.704	0.431	1.200	-0.147
"	Urban,	Single Track	-5.345	0.405	1.039	-0.115
"	"	Multiple Tracks	-3.826	0.295	0.550	-0.034
Flashers,	Rural,	Single Track	-6.160	0.495	0.821	-0.080
"	"	Multiple Tracks	-5.146	0.392	0.538	-
"	Urban	Single Track	-5.719	0.457	0.780	-0.060
"	"	Multiple Tracks	-4.055	0.350	0.378	-0.011

per day, and has recorded two accidents in the last 10 years. Using Equation 9, the estimate of $E\{m\}$ for such crossings is 0.139 accident in 5 years or 0.278 accident in a 10-year period. On this basis and using Equation 10, the estimate of $VAR\{m\}$ is 0.148 for a 10-year period. Therefore, by Equations 5b and 6, $\hat{m}_{10}$ = 0.876 accident or 0.09 accident per year.

The intent of this section is to describe a logically sound procedure for estimating the expected number of accidents at a grade crossing when something is known about the number expected at similar crossings and also about the accident record of that particular crossing. It turns out that the estimator to use is a linear combination of both ingredients (Equation 5). This is the form that the DOT General Accident Prediction Formula takes. It was suggested by Mengert in 1980 *(3)*. The weight to be attached to each ingredient depends on how variable the *m*'s are within the group of crossings considered similar. Two methods for obtaining numerical estimates of $E\{m\}$ and $VAR\{m\}$ have been described.

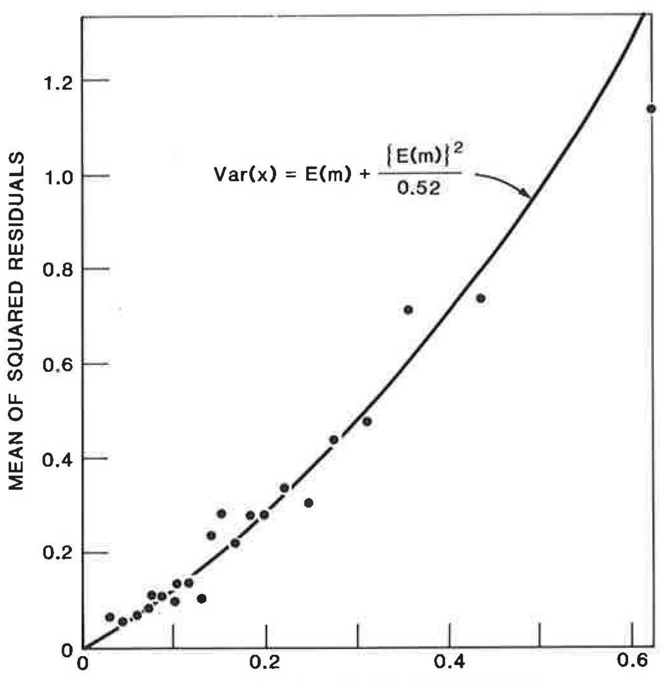


FIGURE 1 Plot of means of squared residuals for groups of crossings with similar estimates of $E\{m\}$.

Data can be used to provide empirical support for results that were obtained earlier by deduction from plausible premises.

Consider, for example, the data in the first two columns of Table 5. These data pertain to the same 9,939 crossings on which the earlier tables are based.

For the 3 years 1980 to 1982, $\bar{x}$ = 0.0478 and s^2 = 0.0570. Thus, if a site recorded two accidents in the 1980 to 1982 period, its *m* is estimated (by Equations 7 and 8, setting $i = j$) to be 0.3629 accidents for that period. Therefore, were there no secular changes in the 5 years 1980 to 1984, 0.3629 × 2/3 = 0.24 accident could be expected from 1980 to 1982. This compares with the observed average of 0.19. To account for the secular trend it might be better to multiply by the ratio of accidents instead of years. The total number of accidents that occurred on these 9,939 crossings was 476, whereas in the subsequent 2 years 262 accidents occurred (see data in Table 1). Therefore, for the 2 years 1983 to 1984, the estimate is 0.3629 × 262/476 = 0.1997 accident. This is the entry in Column 3 of Table 5 against x = 2. In Column 4, the average number of accidents actually observed on these crossings during the 2 years 1983 to 1984 was recorded. If the estimate in Column 3 is any good, there should be a close correspondence between the entries of Columns 3 and 4 in those rows for which the observed average is based on a fairly large number of accidents. This is in fact so. It appears, therefore, that the data on accident occurrence at grade crossings do provide empirical support for the deductions in Equations 5–8.

Early in this paper the question was raised as to whether much is lost by disregarding in estimation either the accident history of a crossing or the characteristic accident experience of similar crossings. It is now possible to answer "yes" to both parts of the question. To see why, imagine that the task is to estimate the *m* for 1980 to 1982 for one of the crossings that recorded two accidents in that period. Recall that the correct estimate is 0.36 accident and that this is confirmed by observations on 36 similar crossings in 1983 and 1984. Had the fact that two accidents were recorded been disregarded, the *m* of the

TABLE 5 PREDICTED AND OBSERVED ACCIDENTS

1	2	3	4
Number of Accidents [x]	Number of Crossings with x Accidents during 1980-1982 [n(x)]	Average No. of Accidents per crossing in 1983-84 Estimated	Observed
0	9512	0.0221	0.0216
1	385	0.1106	0.117
2	36	0.1997	0.19
3	5	0.2876	0.4
4	1	0.3774	3.

NOTE: Urban; single track; 1–2 trains; ADT = 0–1000 vehicles; crossbucks.

crossing would have been estimated to be 476/9,939 = 0.05 during the period 1980 to 1982. The correct estimate is seven times higher. Similarly, had the accident history alone been relied on, the estimate of m would be 2. This overestimates m by a factor of 5.5.

In this section the discussion focused on the question of how to estimate the safety of a crossing. It is concluded that the best estimate is a linear combination of the accident history of the crossing under scrutiny with the mean accident experience of similar crossings. The weight to be used in combining these two sources of information depends on the variance-to-mean ratio of the m's.

These results apply to "steady state" conditions under which the m of a crossing is regarded as constant in time. This is, in general, not the case. The m of every entity can be expected to change in time. How to estimate under such circumstances requires further exploration. The authors have provided only an ad-hoc answer here assuming that each crossing follows the general secular trend.

The results have broad application. The word crossing need only be replaced by the word "intersection," "road section," "driver," or "airport." It follows that the theory and procedures of this section can be applied to two common tasks. The first task is that of identifying deviants, be these unsafe crossings, accident blackspots, or dangerous drivers. For an application to the examination of high-accident road sections see Hauer and Persaud *(8)*. The second task is that of estimating the safety effect of treatment from before-and-after data. The results of several such applications to intersections are discussed elsewhere *(9–11)*. This second task, the estimation of the safety effect of warning devices at crossings, is discussed next.

THE SAFETY EFFECT OF WARNING DEVICES AT CROSSINGS

The principal warning devices used at crossings are crossbucks, flashers, and flashers with gates. The safety effect of replacing a warning device with a costlier one has been investigated in two basic ways. First, by comparing the accident histories of crossings before and after conversion from one type of warning device to another. Such studies are said to be of the before-and-after type. Second, by comparing the safety of crossings that are similar in all measured characteristics except that one crossing is equipped with, for example, crossbucks, whereas the other is equipped with flashers. Such studies are often called cross-section studies. Both approaches have their strong and weak points.

The main results of five major before-and-after studies are given in Table 6.

Estimates of the safety effect given in Table 6 are remarkably consistent. Such consistency is not often found in research on safety and might inspire confidence. Were it not that results of cross-section studies tell a different story, the inclination might be to accept these estimates as correct and lay the matter to rest.

Data are used in a cross-section study to flesh out a model that estimates the expected number of accidents as a function of several causal factors. The DOT Basic Accident Prediction Formula *(1)* is the result of one such effort. Using such models, it is then possible to compare the estimates of the expected number of accidents for two crossings that are identical in all measured causal factors except that they are equipped with different warning devices. Such a comparison is then assumed to reflect the safety effect of the different warning devices when used under identical conditions.

The model published by Schultz et al. in 1969 implies that replacing crossbucks with flashers leaves the hazard index unchanged *(17)*. Eight years later, Coleman and Stewart *(6)* published their results. The model equations appear to indicate that sometimes flashers are better than crossbucks (single-track, rural), sometimes the reverse is true (multiple-track, urban), and sometimes the two types of protection are equal in accident rate (multiple-track, rural). The most recent effort in this direction is the DOT Accident Prediction Formula *(1)*. Under some typical and identical conditions, replacing crossbucks with flashers appears to reduce the probability of accident occurrence by about 30 percent.

TABLE 6 ESTIMATES OF PERCENT REDUCTION IN ACCIDENTS FROM (UNCONTROLLED) BEFORE-AND-AFTER STUDIES

Author:	Calif. P. U. C. (12)	Morris-sey (13)	Coleman (14)	Eck & Halkias (15)	Farr & Hitz (16)
Year Published:	1974	1981	1982	1985	1985
No. of Crossings:	1552	2994	N. A.	7734	5903
Passive to Flashers	64%	65%	71%	69%	70%
Passive to Gates	88%	84%	82%	84%	83%
Flashers to Gates	66%	64%	69%	72%	72%

It appears that uncontrolled before-and-after comparisons indicate that the upgrading of warning devices at grade crossings leads to substantial safety benefits. In contrast, cross-section comparisons lead to an entirely different conclusion. Two questions arise: (a) How can these discrepant findings be explained? and (b) Which conclusion is correct?

When sites are selected for treatment because they appear to be hazardous and when no "matched control sites" are used, estimates of safety effectiveness based on simple before-and-after comparisons are known to be inflated. Because a conversion from passive to active warning is on occasion motivated by the accident history of that crossing, a built-in and systematic bias exists in such before-and-after comparisons. It makes conversion look more effective than it really is. To see why, consider the 143 crossings (Table 1) that had one or more accidents in 1981. During 1981 these crossings recorded a total of 148 accidents. The same crossings (which continued to be equipped with crossbucks) recorded only 16 accidents in 1982. The drop from 148 accidents to 16 cannot be attributed to a change in causal factors; it occurred because in 1981 these crossings had a larger than average number of accidents, and in 1982 they returned to their natural mean. Imagine now what the results of a before-and-after comparison would be had the warning devices at these 143 crossings been upgraded to flashers on January 1, 1982. To attribute the drop in accidents from 148 to 16 (which would have occurred anyway) to the change of warning device would inflate its estimate of safety effectiveness.

The before-and-after studies (the results of which are given in Table 6) did not use matched control crossings to avoid the danger of bias-by-selection. It follows that without a reanalysis of the data, it is impossible to say how much of the effect indicated in Table 6 is illusory and how much of it is genuine.

Similarly, estimates of safety effectiveness that are based on cross-section studies are also in danger of being wrong. To see why, remember that it is under the more difficult circumstances (poor sight distances, downgrades, higher approach speed, more heavy trucks, proximity to schools, etc.) and perhaps because of a history of accident occurrence, that the costlier protection device tends to be used. Therefore, on the average, the costlier the warning device, the higher was the m of the crossing to begin with. This should not be forgotten when the results of cross-section studies are interpreted. When the safety of crossings "with" a warning device is compared with the safety of crossings "without" a warning device, one is comparing "high-m" crossings with the palliative effect of a warning device to "lesser-m" crossings without the warning device. Obviously, the effect of the warning device will appear to be smaller than it really is.

Although it is still not known how to disentangle cause and effect in cross-section studies, methods to cleanse before-and-after comparisons of bias-by-selection are now available *(9)*. It so happens that the methods that provide correct estimates of the expected number of accidents at a crossing are the same methods that facilitate the efficient estimation of the safety effect of warning devices. These are the methods espoused in the earlier section on Theory and Evidence.

To show how the methods for efficient estimation are used to estimate the safety effect of a treatment, it is best to continue the thread of earlier illustrations. Consider a crossing similar to those given in Table 1. From 1980 to 1982, the crossing recorded two accidents. Had there been no change in warning device this crossing could be expected to have, on the average, 0.20 accidents between 1983 and 1984 (see Table 5). This estimate is based on the methods described in the Theory and Evidence section. Let this estimate be entered into a "before conversion" column of a ledger. Assume now that flashers replaced the crossbucks at this crossing on January 1, 1983, and that no accidents were recorded there during the 2 years that followed. A zero is recorded in the "after conversion" column of the ledger. Every crossing converted from crossbucks to flashers adds one line to both columns. The sum of entries in the "before conversion" column indicates the number of accidents that should have been expected had conversions not taken place; the sum of entries in the "after conversion" column indicates what did happen with the new warning devices in place. The difference in the two sums allows the safety effect of such conversions to be estimated.

It is now possible to point out the difference between the method used to obtain the estimates given in Table 6 and the method suggested here. The essence of all procedures for the estimation of the safety effect of some measure is a comparison between what would have happened without the measure and what did happen with the measure in place. In the studies given in Table 6, the assumption has been made that, if nothing is changed, the accident history before conversion is a good indication of what "would have happened" in the after period without the conversion. This assumption is in general untrue and has been shown so by extensive empirical evidence similar to that given in Table 5 [see also Hauer *(4)* and Hauer and Persaud *(18)*]. The method suggested here is based on the assumption that, if nothing has changed, the estimator in Equations 5 and 6 gives an estimate of the *m* for the before period, which is the best estimate of what "would have happened" in the after period.

The estimate of *m* suggested in the Theory and Evidence section appears to be a sound and sensible way to merge information about the causal factors that characterize a crossing and its accident history. However, the method described in that section does not account for factors that are unmeasured. Nor does it account for the effect of the engineering judgment that might be used to decide which of several candidate crossings is ultimately selected for improvement. To the extent that the exercise of such judgment results in the selection of a subset of crossings that is materially different from the set of crossings from which the selection has been made, the accuracy of estimation by the suggested method may be impaired. Whether the subset of sites is in fact materially different and, if so, the inaccuracies in estimation that may result, is at present unknown. Nevertheless, it is important to note that the unbiased estimates discussed next do not take into account the possible effects of engineering judgment.

A comprehensive study to revise the currently used estimates of the safety effect of warning devices at crossings (such as those on which Table 6 is based) is now in progress. Preliminary results are given in Table 7. The first three rows in Table 7 give an impression of the extent of the data on which estimates are based. Row 4 provides information similar to that used in biased before-and-after comparisons. The suspicion was that this is not a correct estimate of the expected number of accidents at the converted sites; that it is an overestimate because the occurrence of an accident on a crossing increases the chance that subsequently it will be equipped with a higher level warning device. Rows 5a and 5b give two estimates of the number of accidents expected to occur in the after period had the warning device not been changed. The estimate in parentheses (Row 5a) has been derived by the method referred to as "the first option" in the Theory and Evidence section. The estimate in brackets (Row 5b) has been obtained by multivariate modeling—the second option described in the Theory and Evidence section—by using more homogeneous groups than those used in earlier illustrations. As is clear from the comparisons of entries in Rows 4 and 5, the suspicions noted previously were well founded. The entries in Row 4 are inflated

TABLE 7 ESTIMATES OF THE SAFETY EFFECT OF WARNING DEVICES

	Crossbucks to Flashers	Crossbucks to Gates	Flashers to Gates
1. No. of crossings converted (1981-1983)	891	1037	934
2. No. of "Before" crossing-years	1734	1962	1855
3. No. of "After" crossing-years	1828	2186	1881
4. No. of "Before" accidents*	165.0	239.1	285.7
5a. Expected No. of "After" acc., option I**	(99.4)	(150.8)	(202.1)
5b. Expected No. of "After" acc., option II	[100.8]	[162.0]	[208.0]
6. No. of "After" accidents	49	50	114
7. Apparent reduction (4)-(6)	116.0	189.1	171.7
8. Apparent % reduction [(4)-(6)]/(4)	70%	79%	60%
9a. Unbiased reduction (5a)-(6), option I	(50.4)	(100.8)	(88.1)
9b. Unbiased reduction (5b)-(6), option II	[51.8]	[112.0]	[94.0]
10a. Unbiased % reduction I, [(5a)-(6)]/(5a)	(51%)	(67%)	(44%)
10b. Unbiased % reduction II, [(5b)-(6)]/(5b)	[51%]	[69%]	[45%]

* Corrected to equalize before and after crossing years.

**Estimated number of accidents expected during the "After" crossing-years had the warning device not changed. Corrected for secular trend, but not corrected for changes in train and car traffic.

by "selection bias" and cannot be used to estimate the safety effect of warning devices. The entries in Rows 5a and 5b estimate the same quantity by using the same data.

Although the two estimates are sufficiently similar, an even closer correspondence would apply were it possible to use more homogeneous groups of entities for "option one" estimates (Row 5a). Thus, more credence should be given to the estimates in brackets, which were obtained by multivariate modeling. The difference between the apparent (inflated, biased) safety effect and the actual safety effect is evident when Rows 7 and 8 are compared with Rows 9 and 10.

When crossbucks are changed to flashers, the apparent reduction is from 165 accidents to 49. Actually, much of this reduction would have occurred even if the warning device was not changed. Only the reduction from 100.8 to 49 can be attributed to the safety effect of the conversion. The difference between the apparent and the real reduction is sizeable. Thinking in terms of "accidents saved," the saving is 52, not 116 accidents.

It follows, also, that when crossbucks are converted to flashers, the percent reduction in accidents is not 70 percent (see Row 8) but only 51 percent (see Rows 9a and 9b). Results for the other two types of conversion are 69 percent, not 79 percent, and 45 percent, not 60 percent. The biased estimates of effectiveness in Row 8 (70 percent for conversion from crossbucks to flashers, 79 percent for conversion from crossbucks to gates, and 60 percent for adding gates to flashers) are similar to estimates obtained in the earlier before-and-after studies given in Table 7. This serves to confirm the earlier assertion that the entries in Table 7 are inflated. Were bias-by-selection purged from the earlier studies, results similar to those in Row 10 would apply. The end result is that the provisional correct estimates of safety effect that are cleansed of bias are (a) crossbucks to flashers, 51 percent; (b) crossbucks to gates, 69 percent; and (c) flashers to gates, 45 percent.

It is tempting to check whether the effect of converting crossbucks to flashers and later flashers to gates adds up to the effect obtained by changing from crossbucks to gates in one step. To see that it does, consider the following argument: if, for a group of crossings all equipped with crossbucks, the expected number of accidents is 100, conversion to flashers is estimated to reduce the number of accidents to 49. A further change from flashers to gates is expected to prevent another 22 accidents on the average. Thus, the joint effect is approximately a 73 percent reduction whereas 69 percent was estimated independently.

SUMMARY

The question of "how to join data about the accident history of a site with information about its geometry, traffic, and other characteristics in order to estimate safety" has been posed. It turns out that a coherent and simple estimate is a linear combination of the mean of the m's that characterizes the population of "similar" sites ($E\{m\}$) with the accident count recorded at the site. The proportions in which these two constituent elements are to be mixed depend on the variance-to-mean ratio $VAR\{m\}/E\{m\}$. The application of this theoretical result is illustrated.

To make use of the suggested estimator requires estimates of $E\{m\}$ and $VAR\{m\}$. Two ways in which such estimates can be obtained are presented.

Presently used estimates of the safety effect of warning devices at grade crossings have been derived from a sequence of uncontrolled before-and-after studies. They are inconsistent with what is found by multivariate modeling, and there is reason to believe that they are inflated as a result of bias-by-selection. The method of safety estimation described in this paper has been used to obtain revised estimates of the safety effect of warning devices. The preliminary findings are that conversions (affected during 1981–1983) from crossbucks to flashers, from crossbucks to gates, and from flashers to gates reduced the chance of an accident by 51, 69, and 45 percent, respectively.

ACKNOWLEDGMENT

Work on this subject has been made possible by support from the National Science and Engineering Research Council of Canada, Transport Canada (Road Safety), and the Federal Highway Administration, U.S. Department of Transportation. The data for the analysis has been provided by the Federal Railroad Administration.

REFERENCES

1. E. H. Farr. *Rail-Highway Crossing Resource Allocation Model.* Report FRA-RRS-81-001. Transportation Systems Center, FRA, U.S. Department of Transportation, Cambridge, Mass., April 1981.
2. J. Hitz. *Rail-Highway Crossing Resource Allocation Procedure—User's Guide*, 2nd ed. Report FHWA-1P-86-11, DOT-TSC-FHWA-85-2. U.S. Department of Transportation, July 1986.
3. P. Mengert. *Rail-Highway Crossing Hazard Prediction Results.* Report FRA-RRS-80-02, PB 80 170749. Transportation Systems Center, U.S. Department of Transportation, Cambridge, Mass., 1980.
4. E. Hauer. On the Estimation of the Expected Number of Accidents. *Accident Analysis and Prevention*, Vol. 18, No. 1, 1986, pp. 1–12.
5. R. J. Baker and J. A. Nelder. *The GLIM System*, Release 3, Manual, Rothhamsted Experimental Station, Harpenden, England, 1978.
6. J. Coleman and G. R. Stewart. Investigation of Accident Data for Railroad-Highway Grade Crossings. In *Transportation Research Record 611*, TRB, National Research Council, Washington, D.C., 1976, pp. 60–67.
7. T. P. Hutchinson and A. J. Mayne. The Year-to-Year Variability in the Number of Road Accidents. *Traffic Engineering and Control*, Vol. 18, No. 9, Sept. 1977.
8. E. Hauer and B. N. Persaud. The Problem of Identifying Hazardous Locations Using Accident Data. In *Transportation Research Record 975*, TRB, National Research Council, Washington, D.C., 1984, pp. 36–42.
9. E. Hauer, J. Lovell, and B. N. Persaud. *New Directions for Learning About Safety Effectiveness.* Report FHWA/RD-86-015. FHWA, U.S. Department of Transportation, Jan. 1986.
10. J. Lovell and E. Hauer. The Safety Effect of Conversion to All-Way Stop Control. In *Transportation Research Record 1068*, TRB, National Research Council, Washington, D.C., 1986, pp. 103–107.
11. E. Hauer and J. Lovell. New Directions for Learning About the Safety Effect of Measures. In *Transportation Research Record 1068*, TRB, National Research Council, Washington, D.C., 1986, pp. 96–102.
12. *The Effectiveness of Automatic Protection in Reducing Accident Frequency and Severity at Public Grade Crossings in California.*

California Public Utilities Commission, San Francisco, 1974. (Reprinted FHWA, U.S. Department of Transportation, 1975.)
13. J. Morrissey. *The Effectiveness of Flashing Lights and Flashing Lights with Gates in Reducing Accident Frequency at Public Rail-Highway Crossings, 1975–1978.* Report FRA-RRS-80-005, FRA, U.S. Department of Transportation, Jan. 1981.
14. J. J. Coleman. Railroad Grade Crossing Research Activities. *Proc., 1982 National Conference on Railroad/Highway Safety.* Kansas City, Mo. H. A. Richards (ed.), 1982, pp. 1-B–5-B.
15. R. W. Eck and J. A. Halkias. Further Investigation of the Effectiveness of Warning Devices at Rail-Highway Grade Crossings. In *Transportation Research Record 1010*, TRB, National Research Council, Washington, D.C., 1985, pp. 94–101.
16. E. H. Farr and J. S. Hitz. *Effectiveness of Motorist Warning Devices at Rail-Highway Grade Crossings.* Report FHWA/RD-85/015. FHWA, U.S. Department of Transportation, 1985.
17. T. G. Schultz, W. D. Berg, and J. C. Oppenlander. Evaluation of Rail-Highway Grade Crossing Protection in Rural Areas. In *Highway Research Record 272*, HRB, National Research Council, Washington, D.C., 1969, pp. 14–23.
18. E. Hauer and B. N. Persaud. A Common Bias in Before and After Accident Comparisons and its Elimination. In *Transportation Research Record 905*, TRB, National Research Council, Washington, D.C., 1983, pp. 164–174.

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

Analysis of Railroad-Highway Crossing Active Advance Warning Devices

BRIAN L. BOWMAN

The purpose of this study was to determine which one of three candidate active advance warning devices for use on roadway approaches to rail-highway crossings was the most effective. Each of the candidate devices, developed during a previous study, consisted of a primary message sign, a supplementary, WATCH FOR TRAINS message plate, and two 8-in. amber, alternately flashing beacons. The devices differed only in the configuration and message of the primary sign. The study was conducted at four sites where sight restrictions on the approach resulted in an insufficient safe stopping distance. The train detection circuitry at each site was modified to provide train activation of each advance warning device approximately 10 sec before the activation of the at-grade warning system. Each test device was installed at all four sites. The results of the speed profile analysis during the activated state indicated that the alternately flashing beacons produce a significant decrease in vehicle velocity. Similar analysis, during the unactivated state, revealed that there was no significant difference in vehicle velocities resulting from the use of different primary signs. These results indicate that the test configuration that used a 48-in. standard (W10-1) railroad advance warning sign would be effective in providing motorists the required advance warning.

The number of annual rail-highway crossing accidents has decreased since the maintenance of records was required by the Accident Reports Act of 1910 *(1)*. Statistics are currently available on grade crossing accidents beginning in 1920. During that year, 1,791 persons died in accidents occurring at public grade crossings *(2)*. By comparison, there were only 542 fatalities in 1983 *(2, 3)*. This reduction in accidents is primarily the result of improvements to the railroad crossing environment. These improvements include the increased use of active warning devices (flashing lights and gates), improvements in track circuitry and control logic, and the installation of advance warning signs and pavement markings.

Hazardous crossing environments exist, however, where safety could be further enhanced by the installation of active advance warning devices. Crossing environments in need of these devices are those where sight restrictions on the approach prevent motorists from viewing either the at-grade warning system or a queue of vehicles stopped at the crossing until an insufficient safe stopping distance exists. These devices would become active only on the presence of a train with the purpose of providing motorists sufficient advance warning to permit a safe stop or a reduction in speed. The devices would be intended for use only on approaches to crossings that are equipped with train detection circuitry.

Goodell-Grivas, Inc., 17320 West Eight Mile Road, Southfield, Mich. 48075.

Many roadway jurisdictions have devised and implemented their own active advance warning devices. These devices usually consist of flashing hazard identification beacons in conjunction with standard or unique advance warning signs. The use of these specialized advance warning systems demonstrates an awareness that standard, passive advance warning signs do not provide motorists with adequate warning at certain types of crossings.

BACKGROUND

In the interest of highway safety, the Federal Highway Administration sponsored a project to develop and test prototype active advance warning devices (AAWDs) for use with existing train detection circuitry and associated railroad crossing signals. Completed in the project concentrated on the development of a simple, relatively inexpensive device that would meet several criteria. These criteria were that the device have high conspicuity, a readily understandable and unambiguous message (even in the fail-safe mode), and that it conform with current signing practices. The result was the selection of three candidate advance warning devices consisting of three principal components: (a) a primary message sign with optional directional arrows, (b) a supplemental message plate, and (c) a pair of alternately flashing yellow beacons. All devices used alternately flashing beacons positioned one above and below the primary and supplementary signs as shown in Figure 1. The supplementary message plate, common to the three candidate devices, consisted of a 3-ft × 2-ft (90 × 60 cm) sign with the message WATCH FOR TRAINS. The primary signs identified as candidates by Ruden et al. are described next *(4)*.

- Primary Sign A, shown in Figure 2, was a 48-in. (120-cm) version of the standard-passive warning sign (W10-1) specified in the *Manual on Uniform Traffic Control Devices* (MUTCD). Because of its circular shape and the R X R symbol filling a large portion of the surface area, Primary Sign A was not used with directional arrows.
- Primary Sign B, a diamond-shaped sign with a black legend on a yellow background, incorporated a red X, bracketed by two Rs (R X R). The red X was used to increase the sign's target value. Instead of the X being constructed at 90 degrees, as with the standard W10-1, it was flattened to 60 degrees. The resultant asymetric symbol had the advantage of being 5 to 10 percent longer than the 90 degree X of the W10-1. In addition, the flattened X provided sufficient room for insertion of directional arrows. This sign has a straight arrow option and is shown in Figure 3.

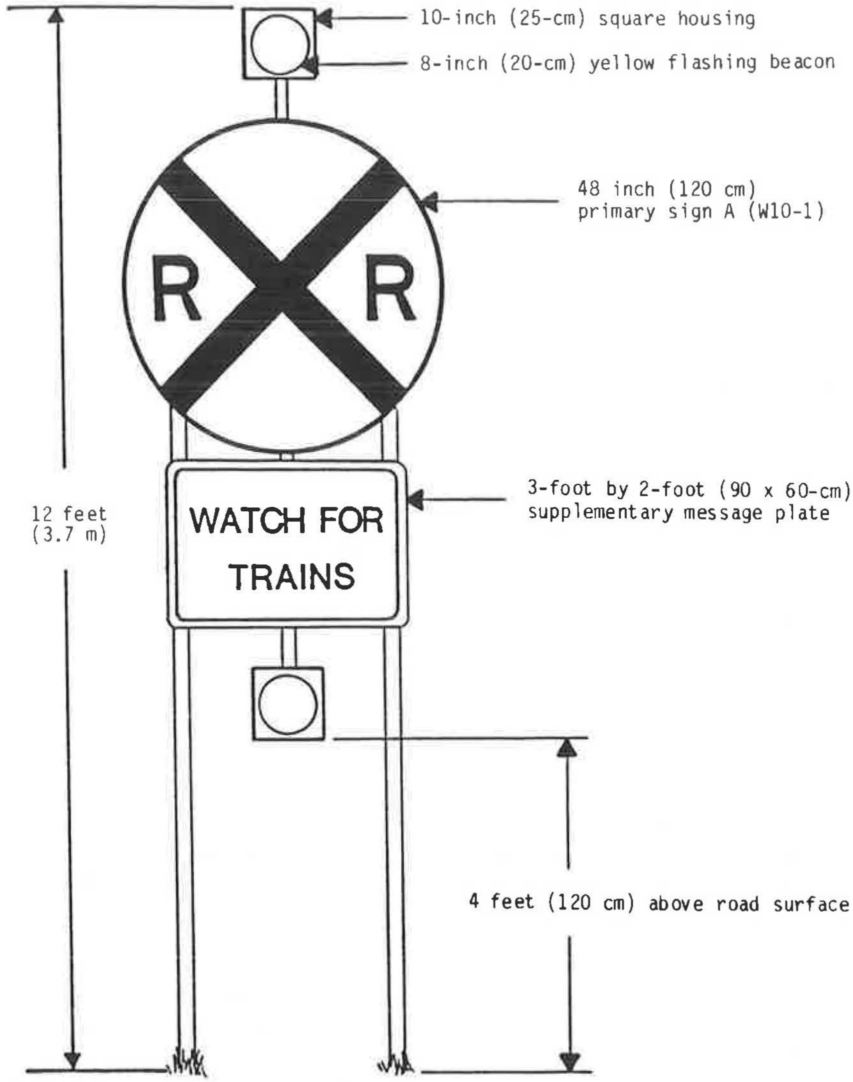


FIGURE 1 Front view of AAWD device.

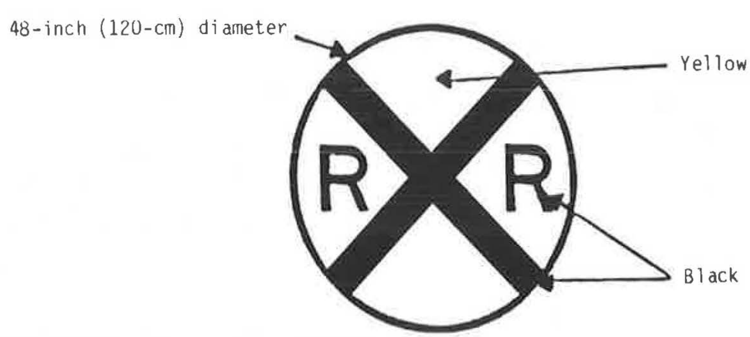


FIGURE 2 Primary Sign A (W10-1).

- Primary Sign C has an arrow option and was intended for use at a horizontal curve (see Figure 4). The sign incorporated a miniature facsimile of the standard W10-1 with red upper and lower quadrants on a yellow background. The miniature W10-1's diameter was one-half the dimensions of the diamond sign.

STUDY SCOPE AND OBJECTIVES

The purpose of this study was to conduct extensive field tests of the three candidate AAWDs to determine the most effective configuration. The effort consisted of three primary tasks: (a) selection of appropriate test sites, (b) modification of the exist-

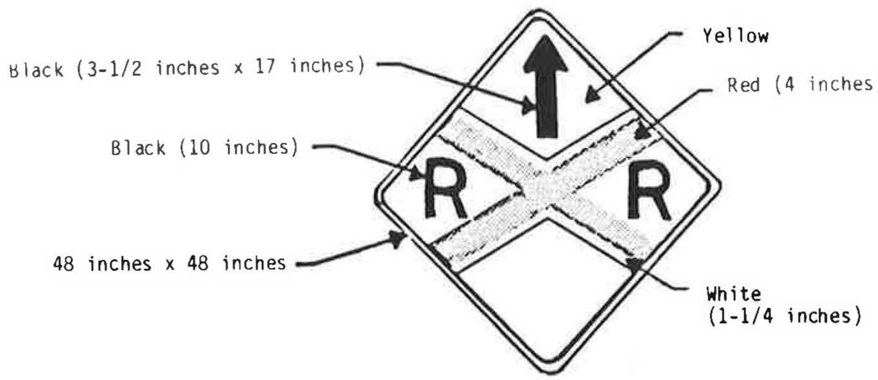


FIGURE 3 Primary Sign B (with vertical curve arrow option).

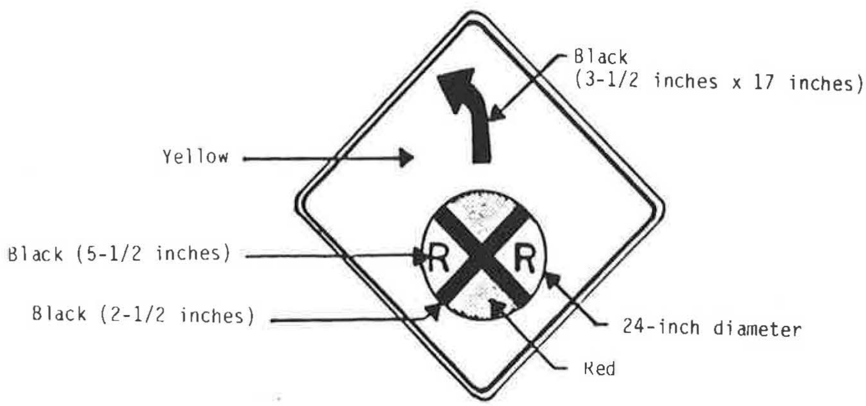


FIGURE 4 Primary Sign C (with left horizontal curve arrow option).

ing train detection circuitry to provide AAWD activation before the start of the crossing signals, and (c) collection of vehicle operational and driver behavior data.

The specific objectives of the study were to

- Perform field demonstration and data collection for each candidate AAWD,
- Analyze the data and evaluate the candidate AAWDs, and
- Determine the most effective AAWD.

METHODOLOGY SUMMARY

Four railroad-highway crossings located in southeastern Michigan were selected as test sites. Each site had (a) sight restrictions that prevented the motorist from observing the crossing on at least one approach, (b) two roadway lanes, (c) a single pair of tracks, and (d) relatively high train and vehicle volumes. The test devices were installed on only one approach to each crossing in the same position as the original advanced warning sign (W10-1). Modifications were made to the train detection circuitry to cause the asynchronous flashing beacons of the test device to activate before the at-grade warning flashers. A summary of selected characteristics for each site is given in Table 1.

Experimental Design

The design used in this project was a modified before-during-after design. It was modified in that the measurements con-

TABLE 1 SUMMARY OF TEST SITE CHARACTERISTICS

Site Designation	ADT	Train Volume	Posted Speed mi/h	Distance from the crossing to AAWD (feet)	Amount of advance activation time (sec).
1	2000	8	45	530	9
2	1100	12	55	560	7
3	6000	10	45	530	9
4	1800	10	45	600	10

1 ft = 0.3 m
1 mi/h = 1.6 km/h

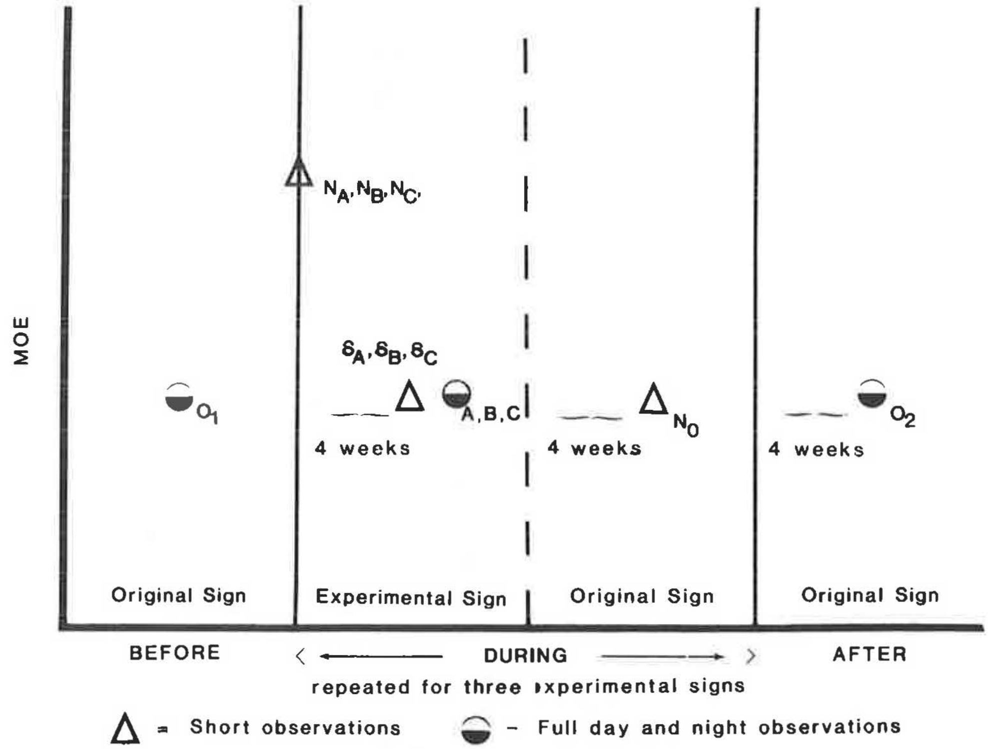


FIGURE 5 Data collection scenario.

tained observations designed to measure the novelty effect of the three different sign installations at each test site. Another deviation from the conventional design is that the during period consisted of three experimental and two intermediate original sign installations. A better understanding of the data collection scenario can be obtained by considering Figure 5 in context with the following paragraphs.

Full day and night observations (O_1) were conducted on the original sign configuration before the installation of any experimental sign. These measures provided the threshold values that were used to gauge the novelty effect and, when analyzed with the after measurements, provided information on long-term trends. On the day that an experimental sign configuration (A, B, or C) was installed, a short observation (denoted by N_i for $i = A, B, C$) was obtained. These observations were interpreted to represent the maximum novelty effect for that particular configuration. The data were statistically analyzed to determine if the measures were significantly different from the appropriate original (O_1) sign observations.

After 4 weeks, another short observation (S_i for $i = A, B, C$) was taken. If the speed observations approached those obtained from the original sign (O_1), then full observations were obtained. If, however, the observations of S_i were similar to N_i, then observation S_i was repeated after 1 week. If the S_i measurements were still similar to N_i then a steady state situation was assumed and full observations were obtained.

After full day and night observations were conducted, the new sign configuration was removed and the original conditions were reestablished. After another 4 weeks had passed, a short reading (N_o) was taken. If this measure was found to be similar to the initial measurements on the original sign configuration (O_1), then the next experimental sign was installed. The same process (denoted by N_B, S_B, B and N_C, S_C, and C) was then repeated for the final sign configuration. After the third experimental sign had been replaced, the original sign was installed and full observations (O_2) were conducted. A flowchart of the data collection procedure is shown in Figure 6.

Evaluation Methodology

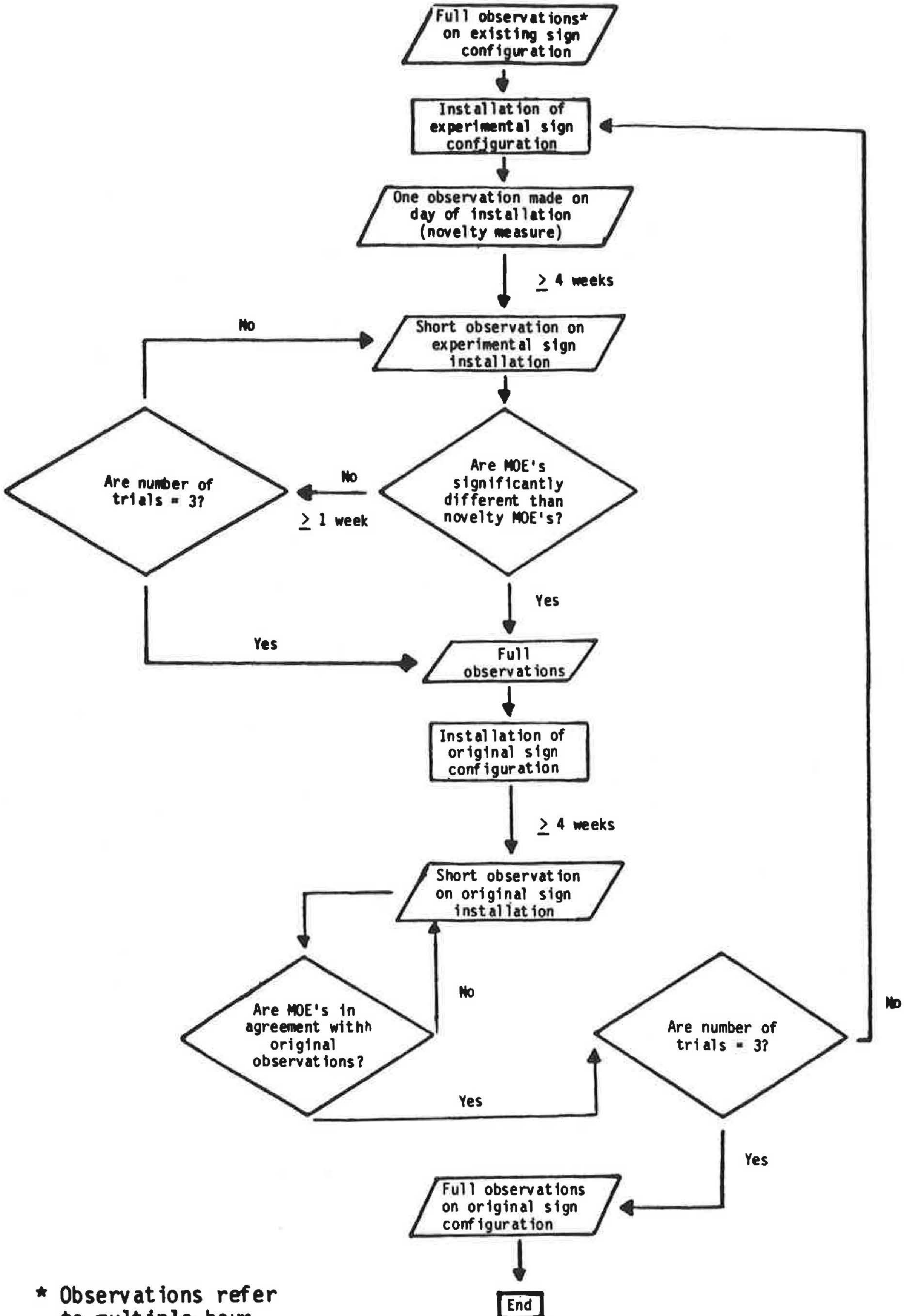
The evaluation methodology consisted of four principal parts: (a) determining the presence of trends over time, (b) ascertaining the presence and dissipation of novelty effects, (c) determining device effectiveness within sites, and (d) determining device effectiveness between sites.

Trends Over Time

The presence of changes resulting from extraneous factors was identified by applying the Scheffe pairwise comparison procedure. This procedure was applied to data that were collected on the original sign in two time periods: before and after the test sign installation procedure. When the Scheffe pairwise comparison procedure resulted in a simultaneous confidence interval that did not include zero, then a significant difference between the before and after data existed. This difference signified a trend, and a pooled mean and variance were calculated. The pooled parameters were then used as the base, or threshold value, for subsequent analysis to determine the effectiveness of the test configurations.

Novelty Effect

An installation and novelty testing procedure was developed that used a 4-week waiting period and statistical tests to ensure



* Observations refer to multiple-hour data collection on each MOE.

FIGURE 6 Flowchart of data collection procedures.

that the novelty effect had dissipated. The statistical tests included plotting a 95 percent confidence interval and performing the Scheffe pairwise comparison procedure to identify significant data differences.

Configuration Effectiveness Within and Between Sites

A two-way analysis of variance (ANOVA) was used to analyze the effects of the different sign configurations on the selected measure of effectiveness (MOE) for each site and between all test sites. ANOVA was used to determine if any exhibited differences were a result of the effect of the devices being tested or a result of the differences exhibited by either the measurement zones or the test sites. If a significant difference was determined as being caused by the test devices, the Scheffe contrast test was used to identify the combination of test devices that caused that difference. When the sample sizes were sufficiently large, separate analyses were performed under the following conditions.

- Daytime/active AAWD
- Nighttime/active AAWD
- Daytime/inactive AAWD
- Nighttime/inactive AAWD

Application of the ANOVA and Scheffe procedures was basically the same for tests of effectiveness both within and between sites. The differences resided in the MOE being analyzed and the structure of the ANOVA matrix. The purpose of the within-site analysis was to determine which test device was the most effective at each particular site. This was performed by testing the mean spot speeds by specific measurement locations at each site against the types of test device.

The purpose of the between-site analysis was to determine the impact of the test configurations irregardless of the test site location. Because this involved testing for the differences between sites, it was necessary to use a MOE that accounted for the various distances of the measurement points from the crossing. The ANOVA for the between-site analysis was performed by testing the average acceleration at every test site against the types of test device.

DATA ANALYSIS RESULTS

Summary of Unactivated Data Analysis

Within-Site Results

The results of the within-site statistical analysis, during the unactivated state, are summarized in Table 2. The data in this table, which result from the site-by-site analysis of the mean spot speeds contained in Tables 3 and 4, indicate that during night conditions at Site 4, Primary Sign A was the only sign that displayed a significant difference from the original sign and every other sign tested. The within-site analysis, therefore, indicated that Primary Sign A was the only sign to display a conclusive impact on vehicle speeds during the unactivated state.

Between-Site Results

The effectiveness of the test devices between sites during the unactivated state was determined by grouping all of the sites together and performing a two-way ANOVA. The purpose of this analysis was to determine if those differences that were identified by analyzing the spot velocities on a site-by-site basis were sufficiently prevalent to result in an overall effect.

Because the magnitude of data variations between the different analysis sites was of interest, it was necessary to provide a measure of effectiveness that was common to all of the sites. Measures based on velocity were not appropriate because the distance from the crossing for the spot speed measurements varied at each site. The larger the distance between the measurement points, the larger the expected velocity change. The overall acceleration was, therefore, used as the measure of effectiveness for comparisons between sites.

The results of the analysis of variance on the overall mean acceleration are given in Tables 5 and 6 for day and night conditions, respectively. The only significant difference revealed by this analysis was that between sites for night conditions, as indicated in Table 6. This result supports those of the within-site analysis, which concluded that the primary signs had similar impacts on vehicle velocities during the unactivated state.

TABLE 2 SUMMARY OF STATISTICAL ANALYSIS OF MEAN SPOT VELOCITIES CONDUCTED WITHIN EACH SITE DURING THE UNACTIVATED STATE

		Site #1				Site #2				Site #3				Site #4			
Condition	Device	O	A	B	C	O	A	B	C	O	A	B	C	O	A	B	C
Day	O	-	-	-	-	-	-	*	-	-	-	-	-	-	-	-	-
	A	-	-	-	-	-	-	*	-	-	-	-	-	-	-	-	-
	B	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Night	O	-	-	*	-	-	-	-	*	-	-	-	-	-	*	-	-
	A	-	-	-	-	-	-	*	*	-	-	-	-	-	-	-	-
	B	-	-	-	-	-	-	-	-	-	*	-	*	-	*	-	-
	C	-	-	*	-	-	-	-	-	-	-	-	-	-	*	-	-

Asterisk (*) indicates significant difference.

TABLE 3 SUMMARY OF DAY MEAN SPOT SPEED (mph) MEASUREMENTS OBTAINED DURING THE UNACTIVATED STATE

Distance from Crossing[1]	Site Designation	Mean Spot Speed (mi/h) for Test Configurations[2]			
		0	A	B	C
1169	1	41.0	44.5	42.8	43.7
1038	2	53.2	51.4	50.6	51.2
1350	3	48.7	49.0	47.4	49.7
845	4	47.3	46.9	48.7	45.8
715	1	41.6	41.1	37.7	40.6
480	2	51.4	50.2	45.4	48.5
545	3	41.9	39.6	41.4	39.1
550	4	39.1	41.9	36.5	35.2
415	1	39.7	31.7	34.9	36.7
240	2	47.7	47.8	44.7	45.2
311	3	41.3	40.8	39.5	39.8
430	4	39.4	38.3	40.1	37.1
215	1	29.8	30.8	28.4	30.3
30	2	40.0	43.1	39.0	42.2
15	3	33.9	30.1	33.6	27.2
270	4	35.0	34.9	35.8	33.7
15	1	26.0	23.4	23.7	23.5
--	2	---	---	---	---
--	3	---	---	---	---
15	4	30.2	32.0	25.4	23.1

1. 1 ft = 0.3 m

2. 1 mi/h = 1.6 km/h

TABLE 4 SUMMARY OF NIGHT MEAN SPOT SPEED (mph) MEASUREMENTS OBTAINED DURING THE UNACTIVATED STATE

Distance from Crossing[1]	Site Designation	Mean Spot Speed (mi/h) for Test Configurations[2]			
		0	A	B	C
1169	1	41.0	42.6	41.6	43.4
1038	2	53.1	54.4	51.3	51.2
1350	3	48.0	48.2	51.8	46.8
845	4	48.1	48.6	49.1	49.4
715	1	41.3	39.4	33.2	39.9
480	2	50.8	51.1	47.3	48.4
545	3	40.5	39.5	46.0	33.7
550	4	36.3	42.9	35.6	35.8
415	1	37.3	34.9	27.9	36.6
240	2	48.9	49.3	46.4	45.3
311	3	40.7	41.4	44.9	36.0
430	4	37.3	39.1	37.2	37.3
215	1	29.9	31.4	28.2	30.9
30	2	44.5	48.4	44.0	41.7
15	3	32.7	30.2	42.0	22.1
270	4	31.4	35.1	29.8	29.8
15	1	25.0	23.5	23.1	23.9
--	2	---	---	---	---
--	3	---	---	---	---
15	4	26.0	30.7	24.2	24.3

1. 1 ft = 0.3 m

2. 1 mi/h = 1.6 km/h

TABLE 5 ANOVA ON THE AVERAGE DAYTIME ACCELERATION (ft/sec^2) FOR EACH DEVICE CONFIGURATION AT THE DIFFERENT TEST SITES DURING THE UNACTIVATED STATE

Test Configuration	Site #1	Site #2	Site #3	Site #4
Original	-1.31	-1.67	-1.01	-1.49
A	-1.42	-1.11	-1.14	-1.69
B	-1.31	-1.09	-1.01	-1.55
C	-1.53	-1.11	-1.36	-1.52

Source	df	SS	MS	$F_{3,9}$	95% Critical F Value
Site	3	0.41	0.14	3.61	3.86
Device	3	0.05	0.02	0.50	3.86
Error	9	0.34	0.04		

TABLE 6 ANOVA ON THE AVERAGE NIGHTTIME ACCELERATION (ft/sec^2) FOR EACH DEVICE CONFIGURATION AT THE DIFFERENT TEST SITES DURING THE UNACTIVATED STATE

Test Configuration	Site #1	Site #2	Site #3	Site #4
Original	-1.31	-1.14	-0.94	-1.74
A	-1.39	-0.63	-1.07	-1.98
B	-0.99	-0.66	-0.73	-1.93
C	-1.47	-1.15	-1.15	-1.98

Source	df	SS	MS	$F_{3,9}$	95% Critical F Value
Site	3	2.55	0.85	28.33*	3.86
Device	3	0.27	0.09	3.00	3.86
Error	9	0.25	0.03		

1 mi/h = 1.6 km/h
Asterisk (*) indicates significance.

Summary of Activated Data Analysis

The number of observations obtained while the devices were activated were, with one exception, too small to perform a statistical analysis. The one exception was when the at-grade flashers and the flashing beacons of Device B were operating in the fail-safe mode because of problems with the train detection circuitry. This provided the opportunity to analyze the differences in the spot velocity measurements at the locations that were influenced by the active advance warning device.

The analysis consisted of comparisons between the mean spot velocities obtained for the activated and unactivated states of Device B at measurement locations that were influenced by the active warning device (S_2, S_3, and S_4). The free speed (S_1) was not used; because of the presence of a horizontal curve, the device was not visible until the vehicles had passed point S_1. The spot velocity at the crossing (S_5) was also not used because it would be expected to be lower for the activated state. Velocities at S_5 would, therefore, include the effect of the at-grade flashers in addition to that of the active warning device.

The activated and unactivated data collected on Device B is given in Table 7. A t-test performed on the area of this table, which is designated by hatch marks, revealed that there was a significant difference between the activated and unactivated data of Device B. This indicates that the flashing beacons are effective in reducing spot velocities.

Summary of Study Results

Analysis of data obtained during the activated condition indicated that the flashing beacons were effective in reducing vehicle speeds. This analysis was performed by concentrating on the spot speeds obtained from locations that were directly influenced by the activated advance warning device. Free-running approach speeds and vehicle speeds obtained at the railroad crossing were not, therefore, included in determining the impact of the devices during the activated state.

With one exception, all of the primary signs evaluated during the unactivated state displayed a similar impact on vehicle velocity and acceleration. The exception was Primary Sign A, which was significantly different from the original sign and all of the other test devices during the night at one site.

An active advance warning device, configured as specified

TABLE 7 ACTIVATED AND UNACTIVATED MEAN SPOT VELOCITY (mph) DATA OBTAINED ON DEVICE B AT 22-mi ROAD

Condition	Test Configuration	Number of Observations	Time of Day	S_1	S_2	S_3	S_4	S_5
Unactivated	B	572	Day	42.8	37.7	34.9	28.4	23.7
Activated	B	61	Day	41.7	27.6	23.7	17.6	4.2

1 mi/h = 1.6 km/h
t-test performed on ////// area
t = 6.2
95-percent critical t value = 2.8

by Ruden et al. with a 48-in. (120-cm) standard (W10-1) railroad advance warning sign (Primary Sign A), would be effective in reducing vehicle speeds during the activated state *(4)*.

COST ANALYSIS

Estimates of the total cost associated with installing an active advance warning device were obtained by itemizing incurred project costs and requesting price quotes from traffic departments located in three different states. These costs are site specific and dependent on the following considerations.

- Costs of train detection circuitry changes. The costs of changes to the train detection circuitry are inherently related to three issues: (a) the economic relationship with the operating railroad, (b) the amount of prior warning time needed, and (c) the amount of at-grade warning time being provided. The first issue is more related to who assumes the cost than to the cost magnitude.

The remaining two issues are closely related. For example, suppose that 12 sec is the amount of time that is desired between the activation of the advance warning device and the start of the grade crossing flashers. Suppose further that the crossing is provided with 40 sec of warning until train arrival, but only 20 sec are required by applicable regulations, existing geometrics, and operating conditions. Under these conditions, there exists an excess of 20 sec from which 12 sec can be provided to the advance warning device. This could result in activating the advance warning device 40 sec and the at-grade flashers 28 sec before train arrival.

When these conditions exist, it is possible to provide the necessary timing changes by installing a capacitive timing relay. This is a relatively inexpensive procedure. When these conditions are not present, it often becomes necessary to extend the detection loop further upstream. This can be expensive, especially if the proximity of adjacent streets complicates the task of extending the loop.

- Electrical connection from crossing control box to active advance warning device. The applicability of providing power to the device by underground trenching or an overhead drop depends on the site environment and preferences of the roadway agency. If overhead wire already exists on the side of the roadway on which the sign is to be installed, then providing an overhead drop is less expensive than trenching and laying conduit. If overhead wiring does not exist, providing overhead capabilities necessitates the installation of a support system such as utility poles. This not only decreases the cost benefits of overhead wiring, but can result in additional roadside hazards.

These considerations were used to develop both low and high installation scenarios. The estimate for an overhead power supply was developed by considering only those installations for which overhead power lines already exist. The high and low installation cost scenarios were used in conjunction with the 1984 accident cost estimates provided by the National Safety Council *(5)*. The results given in Table 8 indicate the number of accidents that need to be prevented to return the installation cost. The components used in developing the high and low cost scenarios are given in Table 9.

INSTALLATION GUIDELINES

Identifying Sites in Need of AAWD Installation

The types of railroad-highway crossings that warrant the installation of an AAWD are characterized as those where the warning devices at the crossing are not visible to vehicles on the approach until an insufficient safe stopping distance exists. One method of initially identifying crossings with sight-restricted approaches is through accident analysis.

Identification through accident analysis requires an investigation of the total number of accidents in the vicinity of the crossing. The accident analysis should be conducted in a manner that is similar to that used for roadway intersections. This involves including all accidents occurring within at least 150-ft (45-m) from the crossing. Approaches with total number of accidents exceeding the areawide mean (for railroad crossing approaches) indicate that further analysis is required. A high incidence of rear-end, run-off-the-road, fixed-object, and train-involved accidents are often an indication of approach sight restrictions.

Ascertaining that a sight restriction contributes to accident occurrence requires an on-site inspection and an application of safe stopping sight distance (SSSD) concepts. The on-site inspection should include obtaining the 85th percentile speed for use in determining the perception-reaction time and the total

TABLE 8 REQUIRED ACCIDENT REDUCTION TO RETURN AAWD INVESTMENT COST FOR DEVICE PLACED AT 600 ft (180 m) FROM CROSSING CONTROL BOX FOR BOTH UNDERGROUND AND OVERHEAD INSTALLATION

		AAWD Costs				Required Accident Reduction			
		Underground		Overhead		Underground		Overhead	
Accident Severity	1984 NSC Costs	Low	High	Low	High	Low	High	Low	High
Fatality	220,000	6,000	10,300	2,000	6,300	1	1	1	1
Personal Injury	9,300	6,000	10,300	2,000	6,300	1	2	1	1
Property Damage	1,190	6,000	10,300	2,000	6,300	5	9	1	1

TABLE 9 ESTIMATE OF ACTIVE ADVANCE WARNING DEVICE FABRICATION AND INSTALLATION COSTS

Activity Sign Fabrication	Itemized Unit Cost[1]		Cost Scenarios Low	High
Square Tube (Pre-galvanized)	2 pcs. @ 39 inches 2 1/2 x 2 1/2 inches 2 pcs. @ 12 inches 2 3/16 x 2 3/16 inches 1 pc. @ 108 inches 1 3/4 x 1 3/4 inches 2 pcs. @ 48 inches 1 3/4 x 1 3/4 inches Fastening hardware Telspar Subtotal	115		
Sign Faces	Primary Sign Supplemental Message	75 30		
Beacons	2 - 8 inch lens units @ $80	160		
Pre-assembly	3 man hours @ $25/hour	75		
Flashing Capability	CD4047 free running multivibrator integrated circuit	100		
	Total Sign Fabrication		555	555
Train Detection Modifications	Pre-emptive method track circuit change	1,000 5,300	1,000	5,300
Installation	Underground (Overhead)		4,400 (450)	4,400 (450)
Approximate totals for underground installation			6,000	10,300
Approximate totals for overhead installation			(2,000)	(6,300)

[1] 1 inch = 2.54 cm.

safe stopping distance. The crossing warning system should be visible to drivers throughout the perception-reaction zone.

If it is determined that insufficient perception-reaction or safe stopping distances exist, then the installation of an active advance warning device may be beneficial. However, consideration should be given to other countermeasures such as additional flashing lights on extended masts, removal of foilage, and other measures to increase the visibility of the crossing warning devices.

Placement Distance From the Crossing

The safe stopping sight distance criteria determines the minimum distance that the AAWD should be placed in advance of the crossing. If necessary, this minimum distance should be increased in order to maximize the distance at which approaching drivers can view the device. For vertical and horizontal curves, this may require that the devices be placed further in advance of the crossing.

Timing of AAWD Activation

The AAWD should be activated before the activation of the crossing warning system by an amount of time equal to the travel time between the AAWD location and the crossing location.

Where a queue of vehicles is expected to occur during the presence of a train, it may be necessary to retain AAWD activation beyond deactivation of the crossing warning system. The amount of retention time will be dependent on the characteristics of each site but can be accomplished by the use of a delay timer.

CONCLUSIONS

The following conclusions were drawn from project activities.

1. During the activated state, the flashing beacons were effective in producing large speed reductions. Statistical analysis performed at the 95 percent level of confidence between the activated and unactivated states revealed (Table 7) that a significant reduction in velocity occurred during the activated state of the device. This velocity reduction occurred in the vicinity of the activated advance warning device.
2. The within-site analysis summarized in Table 2 indicated that Primary Sign A was the only test sign to display a conclusive impact on vehicle speeds during the unactivated state. Primary Sign A [a 48-in. (120-cm) standard (W10-1) railroad advance warning sign] displayed a significant difference from the original sign and all of the other primary signs during the night at one site.
3. The active advance warning device, configured as specified by Ruden et al. *(4)* and using the 48-in. (120-cm) standard (W10-1) railroad advance warning sign, is effective in reducing vehicle approach speed.
4. Practically all of the test configurations, when initially installed, had a novelty effect on the mean velocity of individual vehicles. In most cases, this novelty effect had dissipated after the device was in place for approximately 4 weeks.
5. The approximate cost of device assembly and installation can range from $6,000 to $10,300 for underground and from $2,000 to $6,300 for overhead installation. These costs can be expected to vary from site to site depending on the physical and operational characteristics of the crossing.
6. The most expensive installation scenario would require the prevention of either nine property damage, two personal injury, or one fatal accident during the life of the active warning device to return the investment cost. Accident types to be included in this analysis would include vehicle-train, vehicle-vehicle, fixed object, and run-off-the-road.

REFERENCES

1. *Railroad-Highway Grade Crossing Handbook.* FHWA, U.S. Department of Transportation, Report TS-78-214. Aug. 1978.
2. *Rail-Highway Crossing Safety Report.* FRA, U.S. Department of Transportation, July 1985.
3. *Rail-Highway Crossing Accident/Incident and Inventory Bulletin No. 6.* FRA, U.S. Department of Transportation, June 1984.
4. R. J. Ruden, A. Burg, and J. P. McGuire. *Activated Advance Warning for Railroad Grade Crossings.* Report FHWA/RD-80/003. FHWA, U.S. Department of Transportation, March 1980.
5. *Estimating the Cost of Accidents 1984.* National Safety Council, Chicago, Ill., Aug. 1985.

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

Abridgment

A Comparison of Formulae for Predicting Rail-Highway Crossing Hazards

ARDESHIR FAGHRI AND MICHAEL J. DEMETSKY

The need for improvement at a rail-highway crossing typically is based on the expected accident rate (EAR) in conjunction with other criteria carrying lesser weight. In recent years new models for assessing the need for improvements have been developed, and in the research reported here, five such models selected from a list established from a literature review and a user survey were evaluated. The selected models—the U.S. Department of Transportation (DOT), Peabody-Dimmick, *NCHRP Report 50*, Coleman-Stewart, and New Hampshire—were evaluated using a data base maintained by the Virginia Department of Highways and Transportation. In addition, the performance of the methods for predicting the EAR were compared by using the chi-square test and the power factor. The results indicated that the DOT formula outperformed the other four methods in both the evaluative and comparative analyses, and thus was recommended for use. The priority list produced by this formula is only one criterion used in determining the need to improve conditions at any crossing. This information must be supplemented by regular site inspections and other qualitative issues that cannot be feasibly incorporated into a mathematical formula.

The need for improvement at a rail-highway crossing typically is based on the expected accident rate (EAR) as states use this rate with other criteria to rank crossings. The model used in Virginia to estimate the EAR is documented in *NCHRP Report 50* and is a modified version of the New Hampshire model *(1, 2)*.

Virginia maintains a grade crossing inventory based on the format used by the Federal Highway Administration (FHWA), Federal Railroad Administration (FRA), and the Association of American Railroads (AAR). Part of the information is maintained in a computerized data base, and the remainder is maintained in written form *(1)*. This data base supports the presently used prediction method, but lacks data that some important alternative models require.

In recent years, new methods, such as the U.S. Department of Transportation (DOT) Accident Prediction Formula *(3)* and the Coleman-Stewart model *(4)*, have been developed. With the availability of these methods, the Rail and Public Transportation Division of the Virginia Department of Highways and Transportation requested that several of the most promising methods be evaluated for its use in conjunction with both state and U.S. data bases (DOT, AAR national rail-highway crossing inventory, and FRA accident files). In response a study was conducted to (a) establish a list of nationally recognized models; (b) evaluate representative models for their ability to use available data to show hazard potential at crossings; and (c) recommend whether the currently used method, a modification of it, or a different method should be used by the Rail and Public Transportation Division to predict the accident potential at a crossing.

Virginia Transportation Research Council, Virginia, Department of Highways and Transportation, Charlottesville, Va. 22903.

REVIEW OF AVAILABLE MODELS

Information on 13 nationally recognized models was collected and reviewed *(1)*. These models included the following:

Coleman-Stewart
Peabody-Dimmick
Mississippi
New Hampshire
Ohio
Wisconsin
Contra Costa County
Oregon
North Dakota Rating System
Idaho
Utah
City of Detroit
DOT

The information obtained for seven of these models—the Coleman-Stewart, Peabody-Dimmick, New Hampshire, Oregon, Utah, city of Detroit, and DOT—provided full documentation on their development, testing, verification, and application. In addition to the information collected on these 13 models, data were obtained through a survey questionnaire sent to the departments of transportation in 49 states and the District of Columbia to determine the formulae and methods they use to predict accidents at public rail-highway crossings.

The empirical formulae for calculating hazard indices that have been developed by various organizations and researchers can be categorized into two basic groups. In one group are relative formulae that provide a measure of the relative hazards or the accident expectations at various types of railway crossings. These may be used to rank a large number of crossings in order of priority for improvement, the crossing with the highest hazard index being regarded as potentially the most dangerous and hence the most in need of attention. The second group consists of absolute formulae that forecast the number of accidents likely to occur at a crossing or a number of crossings over a certain time period, and the number of accidents that may be prevented by making improvements at these crossings.

Based on the information obtained and reviewed on the 13 aforementioned models and the results of the survey questionnaire to the states, 5 formulae were selected for testing and evaluation. The DOT, Peabody-Dimmick *(5)*, *NCHRP Report 50*, and Coleman-Stewart represent the absolute formulae. The Coleman-Stewart model, which is relatively new, was included in the evaluation because little is known about its performance. The New Hampshire represented relative formulae.

PRELIMINARY COMPARISONS

The five representative models, though of different forms, share some common features in their basic formats including

1. The use of nationwide data for developing the models.
2. Employment of linear regression techniques for determining the parameters. (Except the DOT model, which was developed by using nonlinear regression analysis.)
3. The expectation that the absolute models cannot predict the exact number of accidents that will occur at a crossing. At best, they can predict only the mean number of expected accidents at a crossing during an extended time period. However, the expected value should be a better indication of the number of accidents that will occur at a location than even that location's history *(2)*.

DATA BASE

The Rail and Public Transportation Division maintains a grade crossing inventory program that was developed by the FHWA, FRA, and the AAR. Part of the information used for predictive purposes is maintained in a computer data base, and the remainder is maintained in written form.

The computer data base is sufficient for computing the New Hampshire, Peabody-Dimmick, and *NCHRP Report 50* models, but must be supplemented to compute the DOT and Coleman-Stewart models. The supplemental data items include the number of through trains per day during daylight hours, maximum timetable speed for each crossing, and highway type. Data on the number of school buses per day per crossing and the sight distance for each crossing were also included to permit further analysis.

For this study, the data base was recorded on an NBI (384k) microcomputer. Three computer programs were written to (a) compute the 5-year accident rate for each crossing according to the four absolute models and the hazard index for the New Hampshire model, (b) perform the chi-square statistical testing for the models, and (c) compute the power factors of the models. The computed numbers of accidents, as well as the hazard index for all the crossings determined by each of the models, were saved on the data diskette.

EVALUATION

Methodology

The two methods described next were used to evaluate the representative models.

1. A statistical chi-square formula of the form

$$\sum_{i=1}^{1,536} \frac{(AO_i - AC_i)^2}{AC_i}$$

was used to determine the relative goodness of fit of the four absolute formulae. In this formula, *AO* is the number of observed accidents, and *AC* is the number of computed accidents for each of the 1,536 crossings. The computed number of accidents according to each of the four representative absolute formulae (DOT, *NCHRP Report 50*, Coleman-Stewart, Peabody-Dimmick) were determined and tested by means of the preceding formula.

2. The primary tool for comparison of the representative relative formula (the New Hampshire model) and the four absolute formulae used in this study is the power factor defined as follows. The 10 percent power factor is the percentage of accidents that occur at the 10 percent most hazardous crossings (as determined by the given hazard index) divided by 10 percent *(6)*. The same type of definition holds for the 5 percent power factor, and so forth. Thus, if $PF(5\%) = 3.0$, then 5 percent of the crossings account for 15 percent ($3 \times 5\% = 15\%$) of the accidents (when the 5 percent referred to is the 5 percent most hazardous according to the hazard index in question).

RESULTS

The chi-square tests on the four absolute models indicated that the number of accidents computed by the basic DOT formula had the closest fit to the actual number of accidents at all of the crossings. The summation of chi squares for all of the crossings by the four absolute models is given in the following table.

Model	*Chi Squares*
Peabody-Dimmick	2175.609
NCHRP Report 50	3810.222
Coleman-Stewart	961.166
DOT	833.096

The performance of all five representative models in the second test (the power factor) is summarized in Table 1.

The data in Table 1 indicate the stability of the basic DOT formula as compared with the other four. Research results have also indicated that once the accident history is incorporated into the basic DOT formula, that is, the main DOT formula is used, the DOT power factors for different percentiles of hazard will be significantly better than those of any other model *(6)*.

Testing the Significance of Other Variables

In order to study the significance and possible inclusion of other important variables in the final hazard prediction formula, data were obtained on 9 crossings that had restricted sight distances and 913 crossings that had school bus traffic.

The nine crossings that had inadequate sight distances were statistically insignificant because the 5-year accident data did not indicate the occurrence of an accident on any of these

TABLE 1 RANKING OF THE REPRESENTATIVE MODELS IN THE POWER FACTOR TEST

Percentage of Crossings	Rank 1	2	3	4	5
1	DOT	New Hampshire	*NCHRP Report 50*	Peabody-Dimmick	Coleman-Stewart
2	DOT	New Hampshire	*NCHRP Report 50*	Peabody-Dimmick	Coleman-Stewart
3	DOT	*NCHRP Report 50*	New Hampshire	Peabody-Dimmick	Coleman-Stewart
6	*NCHRP Report 50*	DOT	Peabody-Dimmick	Coleman-Stewart	New Hampshire
10	New Hampshire	*NCHRP Report 50*	DOT	Peabody-Dimmick	Coleman-Stewart
20	DOT	Peabody-Dimmick	*NCHRP Report 50*	New Hampshire	Coleman-Stewart
40	DOT	Coleman-Stewart	Peabody-Dimmick	*NCHRP Report 50*	New Hampshire

NOTE: No. 1 has the highest power factor, No. 5 has the lowest.

TABLE 2 SCHOOL BUS DATA

No. of Accidents	Frequency: Total No. Crossings	Percent	No. Crossings With School Bus	Percent	Average Percent of School Bus Total Traffic	Range (%)
0	1,392/1,536	90.60	816/1,392	58.6	–	–
1	130/1,536	8.40	91/130	70.0	1.54	0.10–7.14
2	10/1,536	0.65	5/10	50.0	0.74	0.46–0.96
3	4/1,536	0.26	1/4	25.0	1.94	1.94

crossings. A summary of the statistics regarding the school bus traffic on the 913 crossings is given in Table 2.

As can be seen from Table 2, of all the crossings that experienced one accident during the last 5 years, 70 percent had an average of 1.54 percent daily school bus traffic. Fifty percent of all crossings that experienced two accidents had an average of 0.74 percent daily school bus traffic, and 25 percent of the crossings with three accidents had 1.94 percent daily school bus traffic.

It can thus be concluded that the two variables—sight distance and number of school buses—are statistically insignificant, and that their inclusion in the final hazard prediction formula will not alter the results.

CONCLUSIONS

In this study, the DOT accident prediction formula outperformed the other four nationally recognized accident prediction formulae. The DOT formula is fully documented in the *Rail-Highway Resource Allocation Procedure User's Guide*. Also described in the guide is a resource allocation model that, together with the accident prediction formula, provides an automated and systematic means of making a cost-effective allocation of funds among individual crossings and available improvement options. The FRA will run the DOT model for states, if requested, on receiving an updated version of the states' inventory file.

The DOT accident prediction formula takes into account the most important variables that are statistically significant in predicting accidents at rail-highway crossings. However, it must be noted that there is no general consensus as to which of the site characteristics are the most important. As a result, the priority list that is produced by using this formula must serve as only one of the criteria for improving conditions at any crossing. This information must be supplemented by regular site inspections and other qualitative issues that cannot be feasibly incorporated into a mathematical formula.

REFERENCES

1. Wilbur Smith and Associates. *A Study of Rail Freight and Passenger Service Improvement Options.* Rail Transportation Division, Virginia Department of Highways and Transportation, Charlottesville, Oct. 1984.
2. *NCHRP Report 50: Factors Influencing Safety at Highway-Rail Grade Crossings.* HRB, National Research Council, Washington, D.C., 1968.
3. J. Hitz and M. Cross. *Rail-Highway Crossing Resource Allocation Procedure User's Guide.* Report FHWA-IP-82. Transportation Systems Center, U.S. Department of Transportation, Cambridge, Mass., 1982.
4. J. Coleman and G. R. Stewart. Investigation of Railroad-Highway Grade Crossing Accident Data. In *Transportation Research Record 611*, TRB, National Research Council, Washington, D.C., 1976, pp. 60–67.
5. L. E. Peabody and T. B. Dimmick. Accident Hazards at Grade Crossings. *Public Roads*, Vol. 22, No. 6, Aug. 1941, pp. 123–130.
6. P. Mengert. *Rail-Highway Crossing Prediction Research Results.* Transportation Systems Center, U.S. Department of Transportation, Cambridge, Mass., 1980.

DISCUSSION

W. D. BERG

Department of Civil and Environmental Engineering, University of Wisconsin, 2206 Engineering Bldg., Madison, Wis. 53706.

The authors report on a study that was designed to evaluate several rail-highway grade crossing accident prediction and hazard index models with respect to their potential applicability in the state of Virginia. One aspect of the paper that merits discussion and comment is the evaluation of the significance of sight distance as a hazard-influencing variable.

Although the authors report that 9 selected grade crossings having restricted sight distance (out of a total of 1,536 crossings available for study) did not experience any vehicle-train accidents over a 5-year period, this is not a sufficient basis to conclude that sight distance is a statistically insignificant variable which, if incorporated in an accident prediction model, would not alter the results. Prior research has shown otherwise *(1)*. Unfortunately, sight distance data are expensive to collect and therefore are not often available to model developers.

It should also be noted that the influence of sight distance on safety, and thus accident rates, will vary with the nature of other prevailing conditions at the crossing. For example, given that grade crossings equipped with only passive warning devices experience on an average of about one accident every 20 years, then a 5-year accident history (as used by the authors) may be misleading. The crossing that is truly average will not experience an accident in 19 out of every 20 years. Clearly a 5-year sample period could not be expected to yield the actual average rate of 0.05 accidents per year. Rather, a rate of either 0.0 (no accidents in 5 years), or 0.2 (one accident in 5 years) would be observed; neither would be a good estimate of the mean.

The contribution of sight distance to hazards at those high-exposure crossings equipped with automatic warning devices is related to track configuration (number and alignment), as well as the design of the track circuit. The presence of multiple tracks where one train can obscure a second train creates a sight distance problem for which a common countermeasure is the addition of gates. A set of tracks that approach the crossing from a horizontal curve may not afford adequate sight distance to a motorist (especially a trucker) who has stopped because of activated flashing light signals. If the track circuit design speed is significantly greater than the train approach speed, the sight distance problem will be worsened because of diminished credibility caused by an unnecessarily long warning time (during much of which the train may not be visible).

Based on the foregoing observations, there is little basis to support the authors' contention that sight distance is a statistically insignificant hazard-influencing variable and, if included in a hazard prediction model, would not be likely to alter the results of an application of the model. It is hoped that this discussion will stimulate future research into this important aspect of rail-highway grade crossing safety.

REFERENCE

1. W. D. Berg and J. C. Oppenlander. Accident Analysis at Railroad-Highway Grade Crossings in Urban Areas. *Accident Analysis and Prevention*, Vol. 1, 1969, pp. 129–141.

AUTHORS' CLOSURE

The study reported in this paper was performed in response to a request by the Rail and Public Transportation division of the Virginia Department of Transportation. The scope of the work was confined to evaluation of available methods (developed by others) to evaluate hazard potential at rail-highway crossings. The investigation was further limited to use of data currently available from the state of Virginia and the U.S. Department of Transportation (DOT). The emphasis was therefore on practical applications of the methodology.

The models tested were selected as a result of a literature review and a national survey of users. The most widely used approaches did not include sight distance as an explanatory variable. As a result of discussions with the client who recognized the potential effects of sight distance as well as the number of school buses using a crossing, it was decided to investigate the significance of these two data items on accident potential.

Berg's concern that the study dismissed the significance of sight distance as a hazard-influencing variable is unrealistic in view of the scope and constraints on the study. This view is based on the following facts. First, the conclusions stated that the priority list produced by using the formula must serve as only one criterion for improving conditions at any crossing. In the final sentence of the paper, it is explained that this information must be supplemented by regular site inspections and other qualitative issues that cannot be feasibly incorporated into a mathematical formula. It is implicit that sight distance falls into this latter category. In the study, the data were interpreted to indicate that the large majority of crossings had adequate sight distance and that crossings with inadequate sight distance (9 of 1,536) were not represented in the sample. The suggestion of site observations in conjunction with formula ratings provides the opportunity for officials to detect inadequate sight distance and overrule the initial prioritization. In this sense, sight distance is given priority over the other variables.

It is possible that many of the variables used in the models tested have statistically insignificant coefficients; this is also true of the coefficients for sight distance in Berg's 1969 paper. He states in the discussion that prior research *(1)* has shown otherwise (i.e., the inclusion of a sight distance ratio altered the results). This is not shown in his paper; it only includes the sight distance ratio as one of seven explanatory variables.

The real question is, Why did the models developed after 1969 not include sight distance? The literature did not reveal any correlation analysis between sight distance and accidents using a large data base that is common to the applications at hand. If the transportation community feels strongly about this issue, the DOT should sponsor a study to resolve this issue of sight distance once and for all.

REFERENCE

1. W. D. Berg and J. C. Oppenlander. Accident Analysis at Railroad-Highway Grade Crossings in Urban Areas. *Accident Analysis and Prevention*, Vol. 1, 1969, pp. 129–141.

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