TRANSPORTATION RESEARCH RECORD 1010

Traffic Control Devices and Rail-Highway Crossings

TRB
TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL
WASHINGTON, D.C. 1985

Transportation Research Record 1010
Price $15.60
Editor: Scott C. Herman
Compositor: Lucinda Reeder
Layout: Betty L. Hawkins

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 crossings.

(Transportation research record; 1010)
1. Traffic signs and signals–Addresses, essays, lectures.
2. Railroads–Crossings–Addresses, essays, lectures.
I. National Research Council (U.S.). Transportation Research Board. II. Series.
TE7.H5 no. 1010 380.5s 85-15364 [TE228] [338.3'12]
ISBN 0-309-03822-7 ISSN 0361-1981

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.

1610

Contents

1010

Addresses of Authors

Bikowitz, Edward W., Engineering Research and Development Bureau, New York State Department of Transportation, 1220 Washington Avenue, State Campus, Albany, N.Y. 12232

Carr, William P., Washington State Department of Transportation, Transportation Building, Olympia, Wash. 98504

Carter, Everett C., Department of Civil Engineering, University of Maryland, College Park, Md. 20742

Chang, Myung-Soon, Texas Transportation Institute, Texas A&M University, College Station, Tex. 77843

Cynecki, Michael J., City of Phoenix Streets and Traffic Department, 125 East Washington, Phoenix, Ariz. 85004; formerly with Goodell-Grivas, Inc.

Demetsky, Michael J., Virginia Highway and Transportation Research Council, Box 3817 University Station, Charlottesville, Va. 22903

Duncan, D., Traffic Management and Engineering Office, Ontario Ministry of Transportation and Communications, 1201 Wilson Avenue, Downsview, Ontario M3M 1J8, Canada

Eck, Ronald W., Department of Civil Engineering, West Virginia University, Morgantown, W.Va. 26506

Erdman, John W., Kidde Consultants, Inc., 1020 Cromwell Bridge Road, Baltimore, Md. 21204

Halkias, John A., Department of Civil Engineering, Florida Institute of Technology, Melbourne, Fla. 32901

Harmelink, M. D., Traffic Management and Engineering Office, Ontario Ministry of Transportation and Communications, 1201 Wilson Avenue, Downsview, Ontario M3M 1J8, Canada

Hemsley, G., Traffic Management and Engineering Office, Ontario Ministry of Transportation and Communications, 1201 Wilson Avenue, Downsview, Ontario M3M 1J8, Canada

Hwang, Kuo-Ping, Department of Civil Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Va. 24061

Jennings, Barton E., Union Pacific Railroad, 1416 Dodge Street, Omaha, Nebr. 68179; formerly with University of Virginia

Kuhk, R. W., Traffic Management and Engineering Office, Ontario Ministry of Transportation and Communications, 1201 Wilson Avenue, Downsview, Ontario M3M 1J8, Canada

Ligon, Claude M., Public Service Commission of Maryland, American Building, 15th Floor, 231 E. Baltimore Street, Baltimore, Md. 21202; formerly with AMAF Industries, Inc.

Lin, Feng-Bor, Department of Civil and Environmental Engineering, Clarkson University, Potsdam, N.Y. 13676

Maurer, Dean A., Bureau of Bridge and Roadway Technology, Pennsylvania Department of Transportation, Transportation and Safety Building, Harrisburg, Pa. 17120

McGee, Hugh W., Bellomo-McGee, Inc., 410 Pine Street, S.E., Suite 101, Vienna, Va. 22180

McLaughlin, Roberta L., Bellevue Engineering Department, 4415 Greenwood Avenue, N., Seattle, Wash. 98103

Messer, Carroll J., Texas Transportation Institute, Texas A&M University, College Station, Tex. 77843

Penland, Stephen, Washington State Department of Transportation, Transportation Building, Olympia, Wash. 98504

Radwan, A. Essam, Department of Civil Engineering, Arizona State University, Tempe, Ariz. 85287

Ross, Scott P., Engineering Research and Development Bureau, New York State Department of Transportation, 1220 Washington Avenue, State Campus, Albany, N.Y. 12232

Rutherford, G. Scott, Washington State Transportation Center, University of Washington, 136 More Hall, FX-10, Seattle, Wash. 98195

Ryan, Timothy A., Kidde Consultants, Inc., 1020 Cromwell Bridge Road, Baltimore, Md. 21204

Sabra, Ziad A., Department of Civil Engineering, West Virginia University, Morgantown, W.Va. 26506

Schafer, James A., Washington State Department of Transportation, Transportation Building, Olympia, Wash. 98504

Shapiro, Steven, Goodell-Grivas, Inc., 1901 North Fort Meyer Drive, Suite 303, Arlington, Va. 22209; formerly with PRC Voorhees

Stein, Howard, Insurance Institute for Highway Safety, Watergate 600, Suite 300, Washington, D.C. 20037

Tarnoff, Phil, Farradyne Systems, Inc., 13321 New Hampshire Avenue, Suite 110, Silver Spring, Md. 20904; formerly with PRC Voorhees

Titishov, T., Traffic Management and Engineering Office, Ontario Ministry of Transportation and Communications, 1201 Wilson Avenue, Downsview, Ontario M3M 1J8, Canada

von Borstel, Edwin, Seattle Engineering Department, 600 4th Avenue, Seattle, Wash. 98104

Zador, Paul, Insurance Institute for Highway Safety, Watergate 600, Suite 300, Washington, D.C. 20037

Zegeer, Charles V., Goodell-Grivas, Inc., 17320 West Eight Mile Road, Southfield, Mich. 48075

Effect of Signal Timing on Traffic Flow and Crashes at Signalized Intersections

PAUL ZADOR, HOWARD STEIN, STEVEN SHAPIRO, and PHIL TARNOFF

ABSTRACT

The relationship of the timing of traffic signal clearance intervals (yellow phase plus red light in all directions) to traffic flow and crash rates at signalized intersections was examined. Specially designed traffic data-logging devices provided information on the presence and speed of vehicles and the signal timing for 91 signalized intersections throughout the United States. Results showed that intersections with more adequate clearance intervals had substantially fewer rear-end and right-angle crashes than those with less adequate clearance intervals. The observed flow of traffic through the intersections after the onset of yellow was largely unaffected by variation in the lengths of clearance intervals; thus the proportion of drivers exposed to cross-street traffic decreased as the clearance interval lengths increased. Ideally, clearance intervals should be long enough to allow slower traffic approaching the intersection to cross before the cross-street traffic starts. However, for the intersections examined in this study, the group with the highest average crash rate also had the slowest average crossing speed, the widest cross streets, and the shortest and least adequate clearance intervals. Crash increases associated with deficient clearance intervals may be caused by abrupt stops by drivers who are reluctant to traverse wide cross streets with traffic waiting to start up or by vehicles unable to clear the intersection under cross-street red-light protection.

The timing of traffic signal clearance intervals can affect crash rates at signalized intersections. The clearance interval is the period that covers both the yellow signal phase and any subsequent time during which signals for all approaches are red. When clearance intervals are not properly timed, some drivers may be forced to choose between abruptly stopping or loosing the cross-street red-light protection while crossing the intersection. Abrupt stopping can cause rear-end crashes and the loss of cross-street red-light protection can cause right-angle crashes. Loss of cross-street red-light protection occurs because some drivers who do not stop after the onset of the yellow light will clear the intersection only after the cross-street red light ends and the cross-street traffic begins to move into the intersection. These drivers are at increased risk of a collision with cross-street traffic. The proportion of drivers exposed to this risk depends on the proportion of drivers who do not stop and the proportion among them who do not clear the intersection before the cross-street light turns green.

A method for setting the clearance interval to minimize the number of drivers who can neither stop safely nor clear the intersection before the onset of the red light was published by Gazis et al. in 1960 (1). Their work suggests that the proportion of drivers at risk of a crash can be reduced to zero for drivers traveling at similar speeds, with similar perception and reaction times and similar deceleration rates.

The 1982 edition of the Transportation and Traffic Engineering Handbook (2) recommends the use of 10 ft/sec^2 as the threshold value for the deceleration rate in Gazis' timing formula. The higher value of 15 ft/sec^2 contained in the previous edition of the Handbook was reduced because research had shown the higher value to be incompatible with observed driver behavior; that is, drivers typically would not brake hard to stop after the onset of yellow (3-7). Repeated observations of driver response to the onset of yellow demonstrated that less than 10 to 20 percent of all drivers are either able or willing to decelerate at rates in excess of 15 ft/sec^2. The 10 ft/sec^2 threshold allows cars to brake more slowly and results in a longer clearance interval. Studies have also shown that only about 10 to 20 percent of all drivers will disregard the yellow signal and continue through the intersection when deceleration rates less than 10 ft/sec^2 would have been sufficient for stopping. Moreover, lengthening the clearance intervals was not found to increase the percentage of drivers who disregard the yellow light (7).

A 1980 survey of intersections in the southeast reported that about one-half had clearance intervals shorter than those calculated by using the too-high 15-ft/sec^2 deceleration rate recommended by the Handbook at the time of the survey (5). The survey also found that almost none of the intersections were adequately timed compared with intersections in which clearance intervals were based on the more recently recommended lower deceleration rate of 10 ft/sec^2.

Although driver response to the onset of yellow has been extensively researched, the effect on the rate of intersection crashes caused by departures from the recommended signal timing practice has not been systematically assessed. The measurement of this effect was the principal goal of the present study. The other goal was to model driver response to yellow signal light changes at intersections. The results of an investigation of crashes at 91 intersections from eight metropolitan areas throughout the United States are reported herein.

STUDY APPROACH

Data were obtained on police-reported crashes during 1979 and 1980 and on the average daily traffic (ADT) volumes through the intersecting streets for the 91 intersections studied. The physical layouts of the intersections were recorded, and specially designed devices for logging traffic data were used to monitor signal changes, vehicle speeds, and the times vehicles passed a point on the far side of the intersections.

Preliminary data analysis identified six variables related to traffic flow and crash rates at intersections: cross-street width, estimated average crossing time, indirect measures of yellow signal timing, indirect measures of the yellow and all-red phases of signal timing, the ADT for the monitored street, and its ratio to the cross-street ADT. These variables were used jointly to sort the intersections into eight relatively homogeneous clusters through the standard statistical procedure of cluster analysis. The variation in crash rates between the intersection clusters proved to be statistically significant at the conventional 0.05 level. The eight intersection clusters were then ranked on crash rates in an ascending sequence, and neighboring clusters with nonsignificant crash rate differences were merged into five overlapping intersection cluster groups to smooth out the variations in the other variables.

The average values of more than 30 intersection variables were determined for each of the five intersection cluster groups. These variables included nine crash rates based on alternative definitions, descriptions of the physical layout, and signal timing as well as traffic flow measures both just before and just after the onset of yellow. These measures were analyzed, and factors associated with variations in clearance interval lengths, driver responses to the onset of yellow, and crash rates were identified.

DATA COLLECTION

Traffic Data Logging System

Traffic data were collected at the far side of intersections by using the traffic data logging system developed by PRC Voorhees (8). This system includes an arrangement of Leupold-Stevens steel-jacketed coaxial cables and cable transducers for the detection of vehicles and the traffic data logger (TDL) unit for the processing and storing of the signals received from both the cable transducers and the traffic signal power lines.

A typical cable arrangement is shown in Figure 1. Cables C1 and C2 are approximately 3 ft apart and span the width of the street for one direction of traffic. The other two cables, CR and CC, are laid directly adjacent to C2 and only halfway into lanes 1 and 2, respectively. When a wheel crosses a cable, the cable transducer produces a pulse that is recorded by the TDL.

The TDL consists of an internal clock, a microprocessor, and a cassette tape recorder. The TDL was designed to encode and record the time and the source cable for every cable actuation by the number of vehicle axles. For example, a two-axle vehicle traveling in lane 2 would produce six actuations: Both wheels on the first and on the second axle would actuate C1, C2, and CC (but not CR). Thus the record for this vehicle would contain six events (e.g., two actuations for each of the three cables).

The TDL also monitored the power lines to the traffic signal through four separate input channels. The status of the traffic signal was recorded at each cable actuation. Signal phase changes and their times of occurrence were also encoded as independent events.

The actuation data were processed first to represent axles and then to simulate vehicles. Axles were accounted for by matching corresponding actuations of cables C1 and C2. Axle speeds were calculated by dividing the known C1 to C2 distance by the elapsed crossing time, and axle lane position was determined from the actuation pattern of the cables.

Axles were then combined to represent vehicles by an algorithm on the basis of matching lane position and speed criteria and the relative distances between the axles. Subroutines were developed to sort out the records for special cases such as those that were caused by the axle configurations of large trucks or by a vehicle occupying two lanes. Unmatched, isolated axles were retained as single-axle vehicles. Experience with the TDL system indicated that single-axle records resulted when one of the axles of a two-axle vehicle was incorrectly identified from its actuations. Traffic signal timing was also decoded.

The TDL system and the associated software were tested at two sites, one in Richmond, Virginia, and one in Miami, Florida (see Table 1). About 95 percent of all vehicles noted by human observers were detected by the system, and more than three-quarters of those detected were correctly identified. The mean speed, as measured by hand-held radar guns, was 7 percent greater than the mean speed obtained by the TDL system at the Richmond site and 1 percent less than that obtained by the TDL system at the Miami site.

Intersections

The traffic flow and crash results reported in this paper were based on data collected at 91 intersections during 1980 and 1981. Data were collected both during the day and at night, but only the daytime (6:00 a.m. to 8:00 p.m.) traffic observations and crashes were used in this report. An observation period ranged from 1.5 hr to almost a complete day.

Intersections were selected from eight jurisdictions located in different regions of the United States: Chicago, Illinois; Denver, Colorado; Miami, Florida; Montgomery County, Maryland; Richmond, Virginia; San Diego, California; San Jose, California; and White Plains, New York. A summary of the test locations and the applicable yellow signal traffic laws is given in Table 2. The jurisdictions were chosen on the basis of their willingness to cooperate in the study, availability of crash data, and availability of PRC Voorhees personnel. The intersections chosen represent a wide range of intersection parameters, including

1. Average approach speed (35 to 55 mph),
2. Cross-street width (20 to 124 ft),
3. Yellow phase (2.8 to 5.7 sec), and
4. All-red phase (0 to 3.0 sec).

Intersections located within some of the jurisdictions often were similar in one or more design and signal timing characteristics. For example, almost all of the Denver intersections had a 3-sec yellow phase followed by a 2-sec red phase and did not have any left-turn signal phases. All-red phases

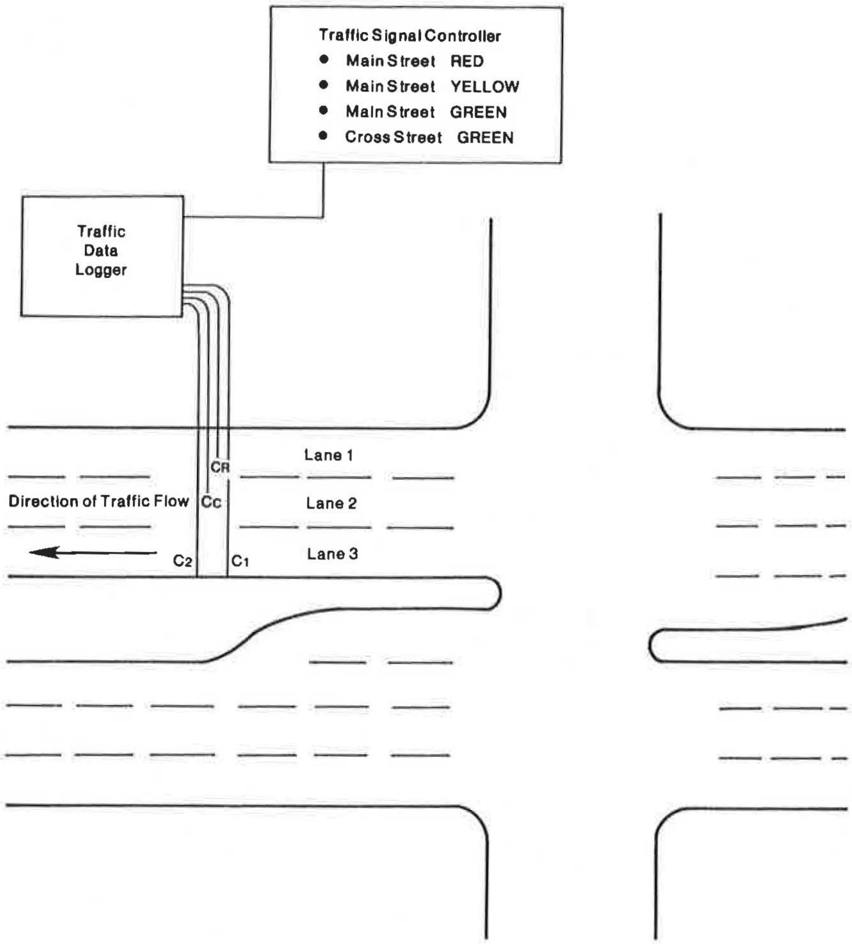


FIGURE 1 Schematic illustration of TDL installation.

were not present or were very short (less than 0.5 sec) at the San Jose sites. Intersections in Richmond, San Diego, and San Jose were typically complex with many independent activated phases. In Miami many of the intersections had permissive left-turn phasing.

Crash Data

Police-reported data were used to identify the intersection crashes during 1979 and 1980 that involved two vehicles. Crashes in which both of the vehicles traveled on the monitored street were grouped together; crashes in which one of the two vehicles traveled on the monitored street and the other on the cross street were placed in a second group. Crashes not fitting either group were not analyzed in the present paper. The shared-approach street crashes of the first group were mostly rear-end crashes. The cross-street crashes of the second group were mostly right-angle crashes.

TABLE 1 Traffic Data Logging System Validation Tests

Test Conditions	Richmond, Virginia	Miami, Florida
No. of lanes	2	3
ADT	23,108	20,000
Speed limit (mph)	45	55
Intersection width (ft)	33	40
Timing (sec)		
Yellow	3.5	5.0
All-red	1.5	1.0
Mean radar speed (mph)	39.6	43.6
Mean TDL speed (mph)	37.2	43.9
TDL performance		
Vehicle detected		
Percent	94	96
No.	312	200
Correctly identified		
Percent	73	81
No.	293	192

DATA ANALYSIS

Six variables, chosen after extensive preliminary analyses of the data, were used for grouping the intersections on the basis of their similarities and

TABLE 2 Intersection Locations and Yellow Signal Laws

Jurisdiction	No. of Intersections	Yellow Signal Law[a]
Chicago, Illinois	2	Stop on yellow
Denver, Colorado	16	Enter on yellow
Miami, Florida	13	Enter on yellow
Montgomery County, Maryland	10	Enter on yellow
Richmond, Virginia	15	Stop on yellow
San Diego, California	11	Enter on yellow
San Jose, California	21	Enter on yellow
White Plains, New York	3	Enter on yellow

[a]The laws were categorized as either allowing an approaching motorist to enter the intersections during yellow or requiring that motorists stop before the intersection if they can safely do so.

differences. The variables were cross-street width (ft), average crossing time (sec), the reciprocal of the braking deceleration rate implied by the yellow phase [DECEL(Y)] (sec^2/ft), the reciprocal of the braking deceleration rate implied by the yellow plus the all-red phase [DECEL(Y+AR)] (sec^2/ft), the ADT on the monitored street (MADT), and its ratio to the cross-street ADT (ADT Ratio).

DECEL(Y) was computed by algebraically solving the Gazis timing formula (1) for the deceleration rate with the observed yellow phase as the clearance interval, and DECEL(Y+AR) was computed the same way from a combination of the yellow and all-red phases. That is, DECEL(Y+AR) = V/{(Y+AR) - t - [(W+L)/V]} in terms of the notation used in the Handbook (2). For some intersections, the yellow phase was so short that it was not sufficient for a vehicle to clear the intersection even if it entered at the beginning of the phase. Thus the estimated value for the deceleration rate based on yellow alone became negative. The resulting numerical instability in the estimate for the deceleration rate would have rendered averages based on it also unstable. The use of the reciprocal deceleration rate circumvented this problem.

A standard cluster analysis package [Ward's algorithm with the STD option (9)] was used to sort intersection data on the six variables into eight disjoint intersection clusters. It was found that cluster membership accounted for about 75 percent of the total intersection variance (R^2 = 0.75) for the six variables.

The rate of crashes per ADT on the monitored street (MADT in 10,000s) was adjusted in proportion to the inverse of the cycle length to allow for the resulting variation in the proportion of vehicles that encountered the onset of a yellow light. The formula used included the average cycle length (about 72 sec) as a scale factor:

ACR = 10,000 x (72/Cycle length)
x (Frequency of shared-approach and cross-street crashes/MADT).

Statistical variability in the adjusted crash rate (ACR) was stabilized by taking square roots, and the variation in $(ACR)^{1/2}$ among the intersection clusters was tested by means of analysis of variance. The relationship between cluster membership and the crash-rate measure [$(ACR)^{1/2}$] was statistically significant ($F_{7,83}$ = 4.36, $p < 0.001$).

The eight disjoint intersection clusters were then ranked according to crash rate estimates in ascending order. Neighboring clusters with crash rates that were similar except for statistical fluctuations were identified by using the Waller-Duncan multiple range test (SAS Institute, 1982), and the intersections in clusters with similar average crash rates were pooled. This procedure yielded five partially overlapping groups of intersections. The first and second groups, for example, included intersection clusters 1 to 4 and 2 to 5, respectively, and overlapped in clusters 2 to 4, but the fifth group included clusters 6 to 8 and overlapped with neither the first nor the second group.

Intersections that were included in cluster groups that did not have significantly different mean adjusted crash rates were pooled. The descriptive statistics for these five groups--labeled A thorough E--are given in Table 3.

As the data in Table 3 indicate, the averages of many of the variables increased or decreased steadily across the five intersection cluster groups. This pattern of variation was further investigated by linearly regressing the variable averages on an index called the relative rank of the intersection cluster group. (By definition, the value of the relative rank increased steadily from cluster group A to cluster group E approximately in proportion to

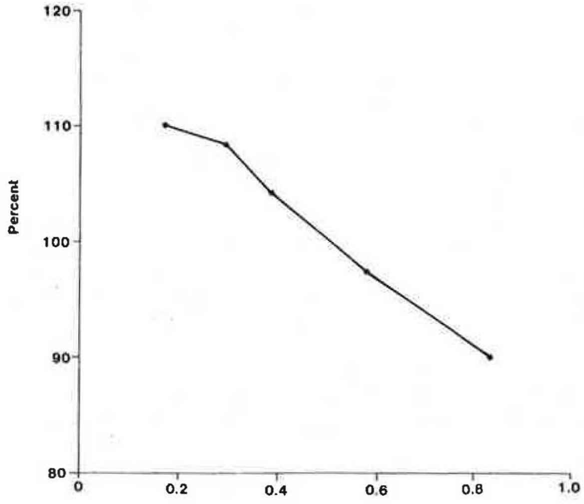
TABLE 3 Intersection Averages for Characteristics by Cluster Group

	Cluster Group Average					Regression on Relative Rank		
Variable[a]	A	B	C	D	E	R^2	Constant	Slope
$(ACR)^{1/2}$	0.92	1.11	1.26	1.53	1.84	1.00	0.70	1.40
$(SACR)^{1/2}$	0.54	0.68	0.78	0.96	1.21	1.00	0.38	1.00
$(CACR)^{1/2}$	0.59	0.71	0.81	1.01	1.18	0.99	0.45	0.91
$(ACR1)^{1/2}$	1.25	1.23	1.23	1.47	1.52	0.84	1.11	0.51
$(SACR1)^{1/2}$	0.79	0.78	0.76	0.89	0.92	0.77	0.71	0.24
$(CACR1)^{1/2}$	0.73	0.76	0.79	0.99	1.03	0.91	0.63	0.51
SCR	1.50	2.18	2.29	2.34	2.73	0.80	1.50	1.55
CCR	1.31	1.84	1.90	2.24	2.24	0.78	1.32	1.30
$(SCR+CCR)^{1/2}$	1.34	1.71	1.80	1.94	2.05	0.83	1.33	0.95
R clearance (sec)	4.76	4.73	4.91	4.96	5.23	0.91	4.58	0.74
Yellow (sec)	4.07	3.73	3.72	3.70	3.88	0.07	3.89	-0.16
All-red (sec)	1.16	1.39	1.37	1.11	0.81	0.60	1.49	-0.70
Clearance ratio	1.10	1.08	1.04	0.97	0.90	0.99	1.16	-0.32
Yellow ratio	0.85	0.79	0.76	0.75	0.74	0.67	0.84	-0.14
All-red ratio	0.24	0.29	0.28	0.22	0.16	0.66	0.32	-0.17
Green phase (sec)	43.3	46.7	45.0	39.7	31.3	0.79	50.8	-21.0
Red phase (sec)	24.5	26.2	36.4	37.0	51.2	0.92	16.9	39.8
Cycle length (sec)	73.0	78.0	86.5	81.6	87.2	0.61	73.1	17.9
DECEL(Y) (sec^2/ft)	0.074	0.062	0.052	0.050	0.042	0.86	0.077	-0.044
DECEL(Y+AR) (sec^2/ft)	0.117	0.115	0.105	0.094	0.076	0.98	0.13	-0.064
F(Y)	1.99	1.53	1.28	1.30	1.18	0.67	1.92	-1.02
F(Y+AR)	2.96	2.72	2.36	2.30	1.89	0.94	3.15	-1.54
RF(Y+AR)	2.74	2.50	2.30	2.47	2.35	0.43	2.67	-0.43
FDIFF	-0.22	-0.22	-0.07	0.17	0.46	0.97	-0.48	1.11
Approach speed (ft/sec)	55.2	53.8	52.4	51.7	48.8	0.97	56.6	-9.2
ADT ratio	4.7	3.1	2.7	2.4	1.3	0.88	4.8	-4.4
Cross street ADT (000s)	8.1	11.7	15.5	14.9	19.4	0.86	6.9	15.3
Monitored street ADT (000s)	21.1	25.3	25.0	21.2	16.6	0.50	26.3	-9.7
Cross-street width (ft)	38.1	39.4	48.7	52.3	67.7	0.95	28.7	45.0
Crossing time (sec)	0.70	0.74	0.98	1.05	1.45	0.95	0.46	1.14
Presence of left-turn lane (%)	12	22	20	19	10	0.16	0.20	-0.08
Relative rank	0.17	0.30	0.39	0.58	0.83	1.00	0.00	1.00

[a]Variable names are explained in text.

the number of intersections included in clusters with lower crash rates.) The relative ranks of the five intersection cluster groups were calculated as follows. First, the intersection clusters were ranked in ascending order from 1 to 8 according to average crash rate. Second, the intersections in the first cluster group were numbered from 1 to 5, those in the second from 6 to 10, those in the third from 11 to 18, and so on until all 91 intersections were numbered. Finally, these numbers were averaged for the intersections in the cluster group and divided by 91 to obtain the relative rank of each.

Because cluster membership accounted for 75 percent of the total intersection variance, and because the cluster groups were formed by pooling clusters with similar crash rates so that crash rate variation within cluster groups was reduced by construction, these linear regressions are likely to give a fairly complete account of all crash-rate-related variation among the intersections. However, because there was considerable overlap among adjacent cluster groups, the R^2 values may overstate the extent to which the relative ranks are linearly related to the other variables. The last three columns in Table 3 present the intercept, slope, and R^2 values for these regressions.



* The relative rank of an intersection cluster group was set to the weighted average of the ranks of the clusters in that group. The individual clusters were initially ranked in ascending order by their crash rates.

FIGURE 2 Average clearance interval as a percentage of recommended value by intersection cluster group.

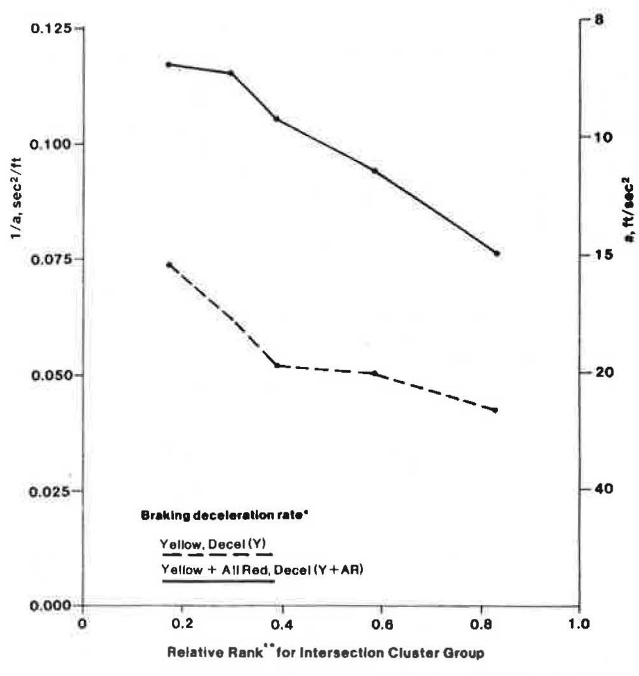
Clearance Interval Averages by Intersection Cluster Groups

Recommended clearance intervals (R clearance in Table 3) were computed for all intersections by using the timing formula with the recommended value of 10 ft/sec^2 for the deceleration rate. In Figure 2 the combined lengths of the observed yellow plus all-red phases were converted to percentages of the recommended clearance intervals (the clearance ratio), averaged within intersection cluster groups, and plotted against the relative ranks for the intersection cluster groups. As Figure 2 shows, these clearance ratios declined steadily across the cluster groups (R^2 = 0.99) from 110 to 90 percent, which indicates that cluster groups with higher relative ranks had less adequate clearance intervals than those with lower relative ranks. The clearance ratios based on the average duration of the yellow phase also declined steadily, from 85 to 74 percent of the total recommended clearance interval, across the intersection cluster groups (R^2 = 0.67).

The relative importance of the monitored streets was measured as the average of the ratio of the ADT on the monitored streets divided by the ADT on the cross streets. The ADT ratio steadily decreased with increasing relative ranks from about 4.2 to about 1.3 (R^2 = 0.88). The data also indicate that relatively more important streets have larger clearance ratios and conversely, relatively less important streets had smaller clearance ratios. Interestingly, the crossing time (vehicle approach speed divided by cross-street width) increased as the clearance interval became shorter. At the opposite extreme, the average monitored street ADT exceeded the average cross-street ADT by the largest amount for the intersection cluster group with clearance intervals closest to the recommended intervals.

Braking Deceleration Rates by Intersection Cluster Groups

The braking deceleration rates implied by the observed lengths of the yellow and combined yellow plus all-red phases were calculated by solving the timing formula algebraically (see Data Analysis section). This solution provides the rate of deceleration that allows a driver traveling at the average intersection speed to either clear the intersection during the yellow phase or stop without entering the intersection. The reciprocals of these rates were averaged within cluster groups; these averages [DECEL(Y) and DECEL(Y+AR)] were plotted against the intersection cluster group relative ranks in Figure 3. As the figure shows, DECEL(Y+AR)



* Recommended clearance interval timing formula was solved in terms of its braking deceleration factor (see text).

** The relative rank of an intersection cluster group was set to the weighted average of the ranks of the clusters in that group. The individual clusters were initially ranked in ascending order by their crash rates.

FIGURE 3 Averages of one over the braking deceleration rates (1/a) implied by the yellow phase and the clearance interval by intersection cluster group.

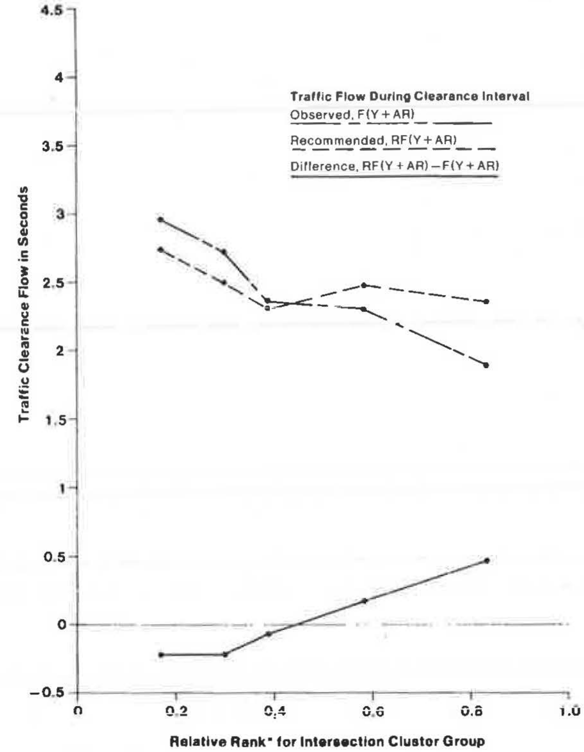
declined steadily between the intersection cluster groups with the lowest and highest relative ranks (R^2 = 0.98). This change in DECEL(Y+AR) corresponds to an increase in the implied braking deceleration rate [1/DECEL(Y+AR)] from 8.5 to 13.2 ft/sec^2. When only the yellow phase is considered, the comparable increase in the braking deceleration rates [1/DECEL(Y)] was from 13.5 to 23.8 ft/sec^2--all higher than the currently recommended rate of 10 ft/sec^2. These results show that at the intersections in cluster groups with high relative ranks, drivers were expected to decelerate at higher rates than at the intersections in cluster groups with low relative ranks.

Average Traffic Flow After Onset of Yellow by Intersection Cluster Group

For each intersection, a base flow rate of vehicles was defined as the average number of vehicles per second that cleared the intersection during 4-sec periods just before the onset of yellow. To assess the response of traffic to the onset of yellow, the average number of vehicles that entered the intersection after the onset of yellow and cleared it during the yellow plus all-red phases was divided by the base flow rate. This ratio, called the total clearance flow [F(Y+AR)], would be proportional to the length of the clearance interval if no vehicles responded to the onset of yellow by stopping. More realistically, the total clearance flow was expected to increase with the length of the clearance interval and to decrease with the proportion of vehicles that stop in response to the onset of yellow.

As the plot of F(Y+AR) against relative ranks of intersection cluster groups in Figure 4 shows, the total clearance flow steadily decreases from about 3 sec for the intersection cluster group with the lowest relative rank to about 1.9 sec for the group with the highest relative rank (R^2 = 0.94). For intersections in the lowest relative rank cluster group, the average number of vehicles that entered the intersection after the onset of yellow and cleared during the clearance interval was about the same as the average number of vehicles that cleared it during the 3 sec just before the onset of yellow. The comparable figure for the highest relative rank cluster group was 1.9 sec. This result shows that as the clearance ratio decreased (see Figure 2), the clearance flow also decreased. If this clearance flow decrease was caused by increased stopping, then the volume of traffic that could have stopped, but did not, at the recommended deceleration rate would also have had to decrease by comparable amounts.

The recommended clearance interval (R clearance) was calculated for each intersection by using the timing formula with a = 10 ft/sec^2. The total clearance flow that corresponds to these standard clearance intervals [RF(Y+AR)] was determined from R clearance in the same way as F(Y+AR) was from the observed clearance intervals. As the plot of RF(Y+AR) in Figure 4 shows, the decrease from 2.7 to 2.4 sec was only about one-third of the comparable decrease in the total clearance flow [F(Y+AR)]. The difference between these quantities [FDIFF = RF(Y+AR) - F(Y+AR)] measures the volume of the flow (in seconds) that failed to clear the intersection during the clearance interval but could have stopped at the recommended 10 ft/sec^2 deceleration. This difference increased steadily with increasing intersection cluster group relative ranks (R^2 = 0.97), which indicates that as the relative lengths of the clearance intervals decreased (see Figure 2), the volume of traffic that failed to clear the intersections during the clearance intervals increased.



* The relative rank of an intersection cluster group was set to the weighted average of the ranks of the clusters in that group. The individual clusters were initially ranked in ascending order by their crash rates.

FIGURE 4 Average traffic clearance flow after the onset of yellow during observed and recommended clearance intervals in multiples of flow rate at time of the onset of yellow by intersection cluster group.

These results suggest that despite variation in the lengths of clearance intervals, driver response to the yellow signal was largely unaffected; consequently, the proportion of drivers who crossed the intersections without protection from cross-street traffic increased when clearance intervals were too short. The slight increase in the flow of traffic traveling during the time period that corresponds to the recommended interval at intersections with deficient or too short clearance intervals may reflect, in part, increased stopping by drivers faced with short clearance intervals and, in part, the enhanced likelihood of drivers responding to a yellow signal of any duration when approaching a cross street with heavy traffic.

Average Crash Rates by Intersection Cluster Groups

The number of police-reported daytime crashes involving two or more vehicles during 1979 and 1980 was divided by the ADT on the monitored street (MADT in 10,000s) and adjusted for cycle frequency per unit of time (see Data Analysis section). The square root of the resulting crash rate [$(ACR)^{1/2}$] was averaged within each intersection cluster group and plotted against the average of the intersection clearance ratios in Figure 5. As the figure shows, $(ACR)^{1/2}$ increased linearly with the clearance ratios, and the highest value of $(ACR)^{1/2}$ was about twice as large as its lowest value. Thus a difference of approximately 20 percent in the ratio of observed to recommended clearance intervals coincided with a difference of a factor of 4 in the ACR.

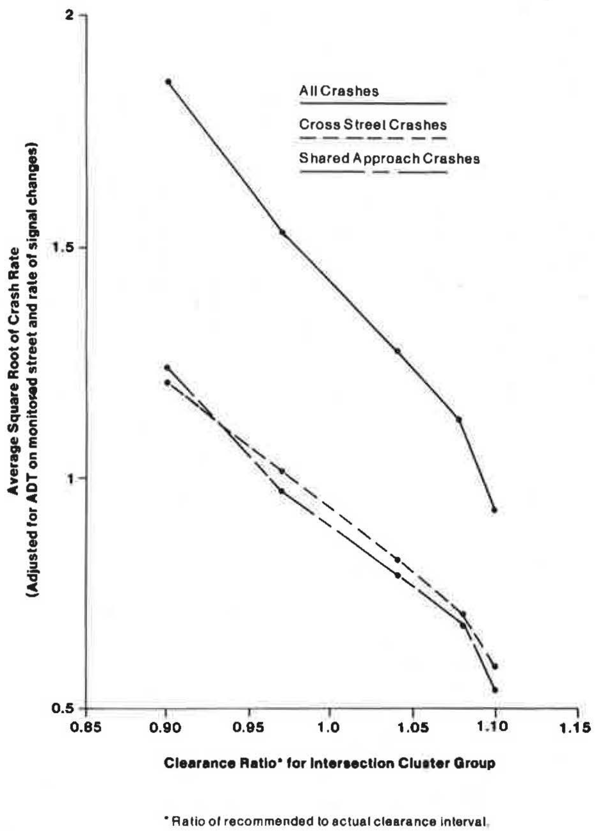


FIGURE 5 Average square root of crash rates by crash type and intersection cluster group.

Similar calculations based on shared-approach crashes alone produced nearly identical results for the adjusted rate of shared-approach crashes (SACR). For cross-street crashes (CACR) that involved one vehicle from the monitored approach and one from the cross street, the results were also nearly identical to those already described.

The sensitivity of these results to the manner of crash rate definition was also explored. Regardless of the manner in which the crash rate was calculated, intersection cluster groups with the least adequate clearance intervals had higher crash rates than those with longer clearance intervals. Even without adjusting for cycle frequencies, intersections with the least adequate clearance intervals had on average 71 percent higher cross-street crashes (CCR) and 82 percent higher shared-approach crashes (SCR) than those with the most adequate clearance intervals. The comparable difference based on the square root of their sum $[(SCR + CCR)^{1/2}]$ was 134 percent.

To allow for the effect of variation in cross-street traffic, the ACR was divided by the cross-street ADT (in 10,000s): ACR1 = 10,000 ACR/ADT. In terms of ACR1, the crash rate of intersection cluster groups with the least adequate clearance intervals exceeded those with the most adequate clearance intervals by about 48 percent. The comparable difference based on cross-street crashes (CACR1) was 99 percent. For shared-approach crashes (SACR1) it was 36 percent.

Miscellaneous Results

The results given in Table 3 show that the five cluster groups differ from one another in almost all respects. Specifically, the average approach speed, the ADT ratio, and the average green phase for the monitored street decreased, whereas the cross-street width, the average crossing time, the cross-street ADT, the red phase, and the complete cycle time increased with increasing relative ranks. The ADT on the monitored street first increased and then decreased.

DISCUSSION OF RESULTS

The relationships between crashes, clearance interval signal timing, and the movement of vehicles reported in this study are based on the analysis of data from 91 signalized intersections in eight metropolitan areas of the United States. Cluster analysis was used to group intersections in terms of their characteristics, the groups were ranked in order of increasing crash rates, and intersections from groups with similar crash rates were combined to form larger groups. Regardless of the manner in which the crash rate was calculated, the intersection groups with the less adequate average clearance intervals had higher average crash rates than those with more adequate average clearance intervals. The combined crash rates for shared approach (e.g., rear-end) and cross-street (e.g., right-angle) crashes differed by 130 percent across the five intersection cluster groups. When these rates were adjusted for signal cycle frequency and ADT on the monitored street, the difference from the lowest to highest crash rates rose to 300 percent. If adjustments for cross-street ADT were also made, the crash rates were still 50 percent greater for the intersections with the least adequate intervals than for those with longer intervals. The variations in crash rates among the cluster groups were associated with specific clearance interval timing, traffic flow, and intersection characteristics.

Crash rates increased as the adequacy of the clearance intervals, based on currently recommended procedures, decreased. The clearance interval durations for the five cluster groups ranged from 10 percent shorter than recommended to 10 percent longer. The crash rate for the group with the least adequate clearance intervals was higher than for the group with the most adequate intervals.

Although the duration of clearance intervals varied across cluster groups, the traffic flow during the clearance interval was largely unaffected. However, clearance interval duration did affect the proportion of drivers who cleared the intersection. The number of drivers who did not clear the intersection during the clearance interval, although they could have stopped at the recommended maximum deceleration rate of 10 ft/sec^2, sharply increased for the intersection groups with the least adequate clearance intervals. Thus, although the traffic flow was similar, the proportion of drivers exposed to cross-street traffic increased and crash rates also increased as the adequacy of the clearance interval decreased.

The ADT on the monitored approach street declined in comparison to the ADT on the cross street as crash rates increased among the cluster groups. As the importance of the monitored street declined, the cross streets were also wider and the monitored traffic slower. These differences resulted in the monitored vehicles requiring increased crossing time to traverse the intersection. If the clearance intervals for these intersections had been calculated on the basis of current recommendations, they would have had longer clearance intervals. These intersections should have had the longest clearance intervals of the intersections studied, whereas they actually had the shortest.

The interpretation of these data by the authors of the overall pattern of association between intersection characteristics, clearance intervals, traffic flow, and crash rates is that the increasing deficiency of clearance interval timing increased the proportion of drivers who would have to stop more quickly than they were accustomed to stopping to avoid entering the intersection without cross-street red protection. However, most drivers cannot or do not stop at high deceleration rates, so that the proportion of drivers who enter intersections and do not clear them during the clearance interval increases sharply. The reduced separation of the two traffic streams and the forced increases in braking lead to substantial increases in crashes.

The most important implication of these results for the practicing traffic engineer is that because drivers cannot be effectively stopped by law from entering intersections after the onset of yellow, it is necessary to time intersections so as to allow a driver who is already in the intersection to clear it before the start-up of the cross-street traffic. Use of the long formula given in the Transportation and Traffic Engineering Handbook (2), with necessary adjustments for the actual path of clearing vehicles, stopline placement, frequency of trucks, and presence of grades, will normally achieve this goal. However, the present analysis was not designed to determine either the optimal signal timing rules or the optimal split between the yellow and all-red phases.

In a recent paper Parsonson and Santiago (5) reviewed a liability suit in which the city of Flint, Michigan, was held responsible for the wrongful death of a driver who died in a crash when his car was hit by a truck at an intersection with an inadequate yellow phase and no all-red phase. The authors of that paper warned the traffic engineering profession that "the traditional design standards for the timing of the clearance period (yellow plus all-red) for traffic signals are inappropriate and unreasonable in some important aspects. They can yield values that are too short for safety...." The authors then recommended improved design procedures "which the engineer would feel more comfortable defending in court."

It has been shown in this paper that even the currently accepted standards are commonly ignored and that clearance intervals that are too short are statistically associated with larger-than-average crash rates. These results and the Flint case should serve to further underline the need to adopt improved clearance interval timing procedures throughout the United States.

ACKNOWLEDGMENTS

The authors would like to thank the traffic engineers and their staffs in the following agencies for their cooperation and assistance in performing the data collection and providing engineering and crash data for the intersections: Metropolitan Dade County Traffic and Transportation Department, Dade County, Florida; Denver Department of Public Works, Denver, Colorado; Illinois Department of Transportation, Chaumburg, Illinois; Kane County Highway Department, Kane County, Illinois; Montgomery County Department of Transportation, Montgomery County, Maryland; San Diego Department of Engineering and Development, San Diego, California; San Jose Department of Public Works, San Jose, California; Virginia Department of Highways and Traffic, Richmond, Virginia; and White Plains Department of Traffic, White Plains, New York.

The authors would like to thank Kenneth Keitt, Marilyn Johnson, and Richard Tanczos of PRC Voorhees for their work in developing the TDL and its software, and in supervising the data collection. The authors would also like to thank Marvin Ginzburg and Sandy Harris of the Insurance Institute for Highway Safety (IIHS) for data processing assistance and Sharon Rasmussen of IIHS for editorial assistance. This work was supported by IIHS.

REFERENCES

1. D. Gazis, R. Herman, and A. Marududin. The Problem of the Amber Signal Light in Traffic Flow. Traffic Engineering, July 1960.
2. L. Rach. Traffic Signals. In Transportation and Traffic Engineering Handbook, 2nd ed., Prentice-Hall, Englewood Cliffs, N.J., 1982, pp. 737-782.
3. P. Olson and R.W. Rothery. Deceleration Levels and Clearance Times Associated with the Amber Phase of Traffic Signals. Traffic Engineering, Vol. 42, No. 4, April 1972, pp. 16-19, 63.
4. W.W. Williams. Driver Behavior During the Yellow Signal Interval. In Transportation Research Record 644, TRB, National Research Council, Washington, D.C., 1977, pp. 75-78.
5. P.S. Parsonson and A. Santiago. Design Standards for Timing the Traffic-Signal Clearance Period Must Be Improved to Avoid Liability. In ITE Compendium of Technical Papers, ITE, Washington, D.C., 1980, pp. 67-71.
6. W.A. Simpson, P.L. Zador, and P.J. Tarnoff. The Influence of the Time Duration of Yellow Traffic Signals on Driver Response. ITE Journal, Vol. 50, No. 11, Nov. 1980, pp. 22-29.
7. P.L. Zador. Driver Behavior at Signalized Intersections in Relation to Yellow Intervals. Insurance Institute for Highway Safety, Washington, D.C., 1980.
8. S. Shapiro, R. Tanczos, and M. Johnson. The Traffic Data Logging System. PRC Voorhees, McLean, Va., Oct. 1982.
9. SAS User's Guide: Statistics. SAS Institute, Cary, N.C., 1972.

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

Right-Turn-on-Red Characteristics and Use of Auxiliary Right-Turn Lanes

FENG-BOR LIN

ABSTRACT

Right-turn-on-red (RTOR) as a means of expediting traffic movements can be complemented with auxiliary right-turn lanes. In this study the characteristics of several RTOR flow parameters are identified from field observations. These and other flow characteristics are analyzed with a simulation model to examine alternative designs and operations relating to RTOR and the use of auxiliary right-turn lanes. To allow an auxiliary right-turn lane to fully serve its function, AASHO recommended that the storage length of such a lane be long enough to prevent a blockage of the traffic in an approach lane. Based on this design requirement, an analytical method for determining the storage lengths is developed.

Intersections are potential bottlenecks in a street network. Searching for ways to improve the operating efficiency of a street network has become an urgent problem because streets are becoming more crowded. Recent efforts in this connection have overwhelmingly focused on the development of better signal control systems. Although this undertaking is necessary, it cannot remove the constraints imposed on traffic movements by the geometric design of an intersection. There are situations where the use of an auxiliary right-turn lane in conjunction with right-turn-on-red (RTOR) can achieve a much needed improvement in the operating efficiency of an intersection.

RTOR is intended primarily as an energy conservation measure. Previous studies (1,2) have indicated that RTOR on urban streets has the potential of reducing fuel consumption by about 5 percent. The implications of RTOR for traffic safety, however, have received more attention in the past (3,4). Several studies (5,6) have also examined the general impact of RTOR on vehicle delays. But current understanding of this subject is still not sufficient to assist traffic engineers in making planning, design, and operating decisions.

The impact of RTOR on the operating efficiency of an intersection can be expected to be small if no auxiliary right-turn lanes are provided. The beneficial effects of an auxiliary right-turn lane can be fully realized only if the lane is sufficiently long. Otherwise, straight-through vehicles arriving during a red phase may block the entrance to the right-turn lane. Similarly, right-turn vehicles may block straight-through vehicles. To prevent such a blockage from occurring, AASHO (7) recommended that the storage length of an auxiliary lane be based on 1.5 to 2 times the average number of vehicles that would be stored per signal cycle. This recommendation is convenient to follow but is vague in the methods needed to obtain an estimate of the average number of stored vehicles per cycle. It may not lead to a proper design in terms of intersection operation or resource allocation. A better method of determining the storage requirements of auxiliary right-turn lanes is needed.

To provide a better understanding of the potential and limitations of using auxiliary right-turn lanes with RTOR to improve intersection operations, several observed RTOR flow characteristics are first described. This is followed by a discussion of the potential impact of auxiliary right-turn lanes with or without RTOR on vehicle delays. Finally, an analytical method for determining required lengths of auxiliary right-turn lanes is presented.

CHARACTERISTICS OF RTOR FLOWS

RTOR flows are associated with a greater variability of driver behavior than other types of directional flows at an intersection. They are also subject to the influence of a greater variety of signal control and traffic flow conditions. As a result, it is difficult to rely entirely on field observations to identify the complex relationships between RTOR and its related variables. An alternative approach is to identify the basic characteristics of RTOR flows from field observations and use them to develop a simulation model as an analysis tool.

For this reason data on various RTOR characteristics were collected in downtown Syracuse, New York, and its suburban area. These data made it possible to quantify several RTOR flow parameters. Included among them are the use of RTOR opportunities, the gap-acceptance behaviors of RTOR drivers, the dwell times of unopposed RTOR vehicles, and the efficiency in executing multiple right-turns-on-red. These parameters are described in the following sections.

Use of RTOR

The data in Table 1 give the proportions of right turns made during red intervals from 16 right lanes with mixed directional flows. The traffic flows at these sites were regulated with signal controls that have cycle lengths that range from 75 to 110 sec. The green phases for the right lanes accounted for 20 to 30 percent of the cycle lengths. The cross-traffic volumes were about 150 to 350 vehicles per hour (vph), with approach speeds less than 35 mph.

Although the red phases accounted for 70 to 80 percent of the cycles, the number of RTOR vehicles as a percentage of the right-turn vehicles was low in most of the lanes examined. In 12 out of the 16 lanes, for example, less than 30 percent of the

TABLE 1 Rates of RTOR

Lane	Flow Rate (vph)	Right-Turn Percentage	Percentage of Right Turns Making RTOR from Traffic Lane	Shoulder
1	150	43	20	7
2	345	63	31	5
3	162	43	29	0
4	162	45	35	3
5	166	15	38	1
6	331	47	41	3
7	376	42	5	36
8	265	35	77	3
9	413	45	17	0
10	272	44	21	0
11	298	54	25	5
12	513	39	11	68
13	431	16	20	0
14	240	22	9	1
15	466	22	11	0
16	135	50	10	4

Note: vph = vehicles per hour.

right turns were made from regular traffic lanes during red phases. The average for the 16 lanes was only 21.8 percent of the right turns.

When the circumstance permitted, a driver made RTOR from a shoulder. It is interesting to note that a small increase in the approach width of an intersection can drastically raise the rate of RTOR use. This phenomenon is exemplified by the RTOR flow rates of Lanes 7 and 12 given in Table 1. Both lanes have an unpaved shoulder area about 6 ft wide. The shoulders are not intended to carry traffic, but a large proportion of the right-turn drivers in either lane pulled onto the shoulder and subsequently executed RTOR. This demonstrates the need to provide exclusive right-turn lanes to accommodate RTOR vehicles.

In contrast, many drivers may elect not to use RTOR opportunities. Based on observations of 359 leading right-turn drivers in 10 right lanes, it was found that the rate of rejection for using RTOR opportunities was 16 percent among the drivers. One major reason for rejecting RTOR is likely to be drivers' ignorance of the RTOR regulation. Current lack of uniformity in RTOR signing may be another factor that contributes to some drivers' reluctance to execute RTOR at certain intersections. The existing rejection rate can be expected to dwindle to a negligible level in the future when drivers become more familiar with RTOR regulation and signing.

Gap-Acceptance Behavior

The rate of RTOR use is also governed in part by the sizes of the gaps (headways) in the cross traffic and by the ability of the right-turn drivers to accept such gaps. To quantify the gap-acceptance behavior of the RTOR drivers, field data on leading right-turn driver movements during red intervals were collected from 10 right lanes.

This task was tedious and difficult mainly because of the lack of suitable intersections where a large sample of data could be obtained in a short period of time. RTOR driver behavior and the existence of mixed directional flows in every lane examined further aggravated the situation. As mentioned previously, some right-turn drivers chose not to execute RTOR even when there was ample opportunity to do so. Other right-turn drivers were blocked by straight-through or left-turn vehicles. And some leading right-turn drivers were able to make RTOR without opposition.

From the 10 right lanes chosen for the study, a total of 359 leading right-turn drivers was observed. Out of this total only 202 rejected at least one gap before merging into the cross traffic. The behavior of these drivers indicates that a gap of less than 5 sec has little chance of being accepted and a gap of greater than 15 sec is unlikely to be rejected. The critical gap of these drivers, as shown in Figure 1, was found to be about 8.4 sec. This gap is considerably longer than the typical critical gaps of 4 to 5.5 sec of opposed left-turn drivers (8,9,10).

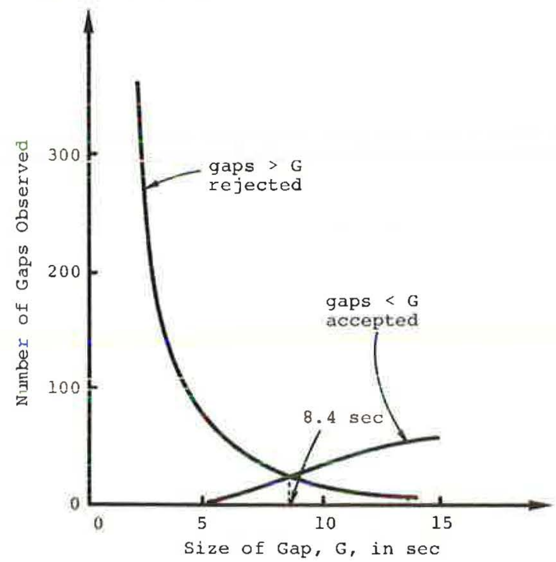


FIGURE 1 Gap-acceptance characteristics of RTOR drivers.

An RTOR driver who accepts a gap will consume a portion of the gap. This is because the turning vehicle has to wait until after the leading cross-traffic vehicle that forms the gap passes the conflicting point. On average, it took the observed RTOR drivers 3.1 sec after the passing of the leading cross-traffic vehicle to execute RTOR. This magnitude of the elapsed time has an adverse impact on the RTOR capacity of a right lane.

Dwell Time of Unopposed RTOR Drivers

RTOR drivers are required to come to a stop before making the turn. This requirement incurs a dwell time for every RTOR driver. The dwell time is defined as the elapsed time from the moment a driver reaches a position from which he can make RTOR until he starts executing the turn. In the field investigation a total of 246 right-turn drivers who made RTOR without any opposition were observed. The average dwell time of these drivers was 4.4 sec. Approximately 40 percent of these drivers executed RTOR within 2 sec of their arrivals at the merging position. This represents the proportion of drivers who violate the requirement to come to a stop.

Multiple RTOR

When a long gap is available in the cross traffic, several drivers may be able to use this gap to execute multiple right-turns-on-red. The time required to do so can affect delays, number of stops, fuel consumption, and the capacity of a right-turn lane. Figure 2 shows the observed relationship of this time requirement to the number of right-turns-on-red made in a gap by a queue of right-turn vehicles. The time requirement was measured from the moment the first RTOR vehicle was in a position to move until the last RTOR vehicle started the turn. The figure shows that two consecutive executions of RTOR would require an average of 10.5 sec. With five multiple turns, the average total time requirement reaches approximately 23.5 sec. This time requirement is long compared with an average of about 14 sec needed for the first five right-turn queuing vehicles to enter an intersection during a green phase. It is obvious that a red phase is only about 60 percent as useful as a green phase of the same length, even when cross traffic does not exist.

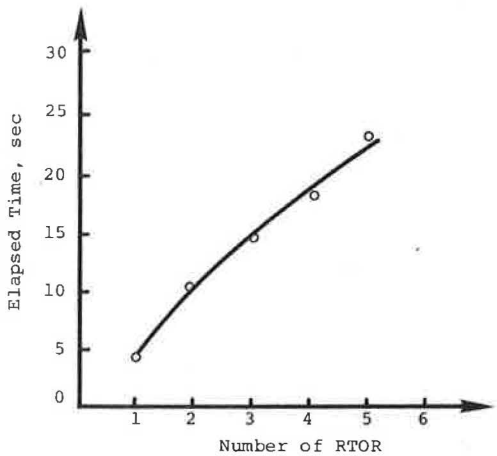


FIGURE 2 Time requirements for multiple RTOR.

EFFECTS OF RIGHT-TURN LANES

The ability of an auxiliary right-turn lane to reduce vehicle delays depends on a number of factors. These factors include the storage length of the right-turn lane, right-turn percentage of the approaching vehicles, flow rates, RTOR policy, type of signal control and signal timing settings, pedestrian flows. The large number of influencing factors precludes a comprehensive analysis of the impact of right-turn lanes. Nevertheless, insight into the potential impact of right-turn lanes can be obtained with an analysis of limited scope. An analysis of this nature is presented herein. The analysis is based on an intersection controlled with a two-phase pretimed signal. Furthermore, pedestrian interferences with the right-turn vehicles are assumed to be negligible and the rightturn lane has a sufficiently long length to avoid blockage of the traffic lanes.

To facilitate the analysis, a simulation model is calibrated in part on the basis of the RTOR data described previously. Field data on straight-through and right-turn queuing flows are also used in the calibration. However, one deviation from the observed RTOR flow characteristics is allowed in the model. This deviation stems from an implicit assumption in the model that every driver will use RTOR opportunities. This assumption could lead to slight overestimates of the impact of RTOR.

The operating efficiency of a signalized intersection is governed to a large extent by the discharge headways of dissipating queuing vehicles. Therefore, such headways are carefully treated in the simulation model. The data in Table 2 give the representative averages of observed queue discharge headways for three types of turning movements. It can be noted from the data in this table that the average discharge headways stabilize at a value of about 2.1 sec for straight-through vehicles and about 2.4 sec for right-turn vehicles. The corresponding saturation flow rates are approximately 1,700 vph for straight-through flows and 1,500 vph for right-turn flows. The variations among individual discharge headways are large.

TABLE 2 Representative Average Queue Discharge Headways

	Avg Discharge Headways (sec)		
Queuing Position	ST	RT	Mixed ST and RT
1	3.3	3.6	3.2
2	2.6	2.8	2.7
3	2.4	2.6	2.5
4	2.3	2.5	2.5
5	2.2	2.5	2.3
6	2.2	2.5	2.2
7	2.1	2.4	2.2
8	2.1	2.4	2.2
9	2.1	2.4	2.2
10	2.1	2.4	2.2

Note: ST = straight through and RT = right turn.

The discharge headways of those vehicles in the same queuing position can be represented in terms of the percentages of their average. With this transformation, it was found that the discharge headways have a distribution that conforms to the one shown in Figure 3. This distribution is applicable to all types of turning movements and to all queuing positions.

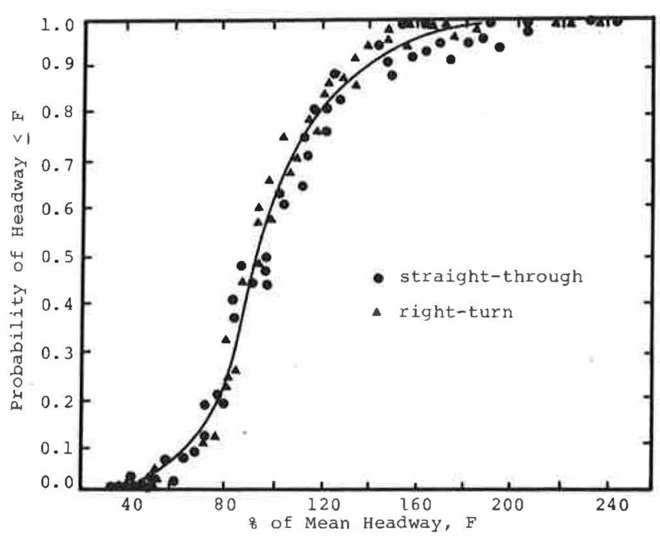


FIGURE 3 Normalized cumulative distribution of headways.

Figure 3 shows that the discharge headways may vary from about 40 percent to more than 240 percent of the averages. The upper bounds of the variations are not the same for all queuing positions. Field data indicate that such upper bounds can be approximated by the following equation:

$$U = 163 + 14.5i + 3i^2 - 0.5i^3 \leq 240 \quad (1)$$

where U is the upper bound and i is the queuing position of a vehicle. The queue discharge characteristics, as represented by Table 2, Figure 3, and Equation 1, are incorporated into the simulation model.

The model is used to examine delays associated with three alternative combinations of geometric design and RTOR policy. The first alternative has an approach lane without an auxiliary right-turn lane and RTOR is not permitted. The second alternative has an approach lane that diverges into a straight-through lane and an auxiliary right-turn lane without RTOR. The last alternative has the same geometric design as the second alternative, but RTOR is allowed.

Without RTOR

Figure 4 shows the average delays of vehicles in a mixed straight-through and right-turn flow under a specific signal control condition. The signal has a cycle length of C = 60 sec and a green phase of G = 26 sec for the arriving vehicles. The figure reveals that the availability of a right-turn lane decreases the average delays by about 20 percent when right-turn vehicles account for 10 percent of the arriving vehicles. When right-turn vehicles account for 40 percent of the arriving vehicles, the delays can be reduced by 20 to 50 percent, depending on the arriving flow rate.

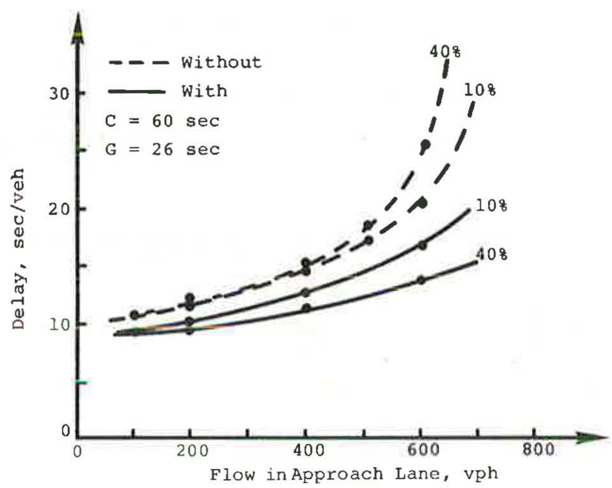


FIGURE 4 Vehicle delays with and without right-turn lanes at two levels of right-turn percentages.

The reductions in delays attributable to the presence of an auxiliary right-turn lane are also affected by signal timing settings. Figure 5 shows that the amount of reduction in delays would usually be less than 5 sec per vehicle under varied flow and signal conditions if the saturation ratio of the approach flow is less than 0.7. The saturation ratio is defined as

$$r = QC/SG_e \quad (2)$$

where

r = saturation ratio;
Q = approach flow rate;
C = cycle length;

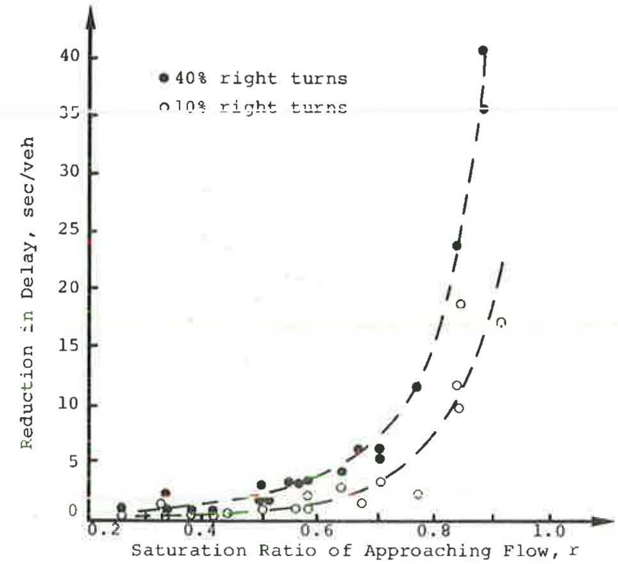


FIGURE 5 Reductions in average delays attributable to auxiliary right-turn lanes without RTOR.

G_e = effective green phase for the arriving vehicles, equal approximately to the green phase (G); and
S = saturation flow rate.

Figure 5 also shows that when the saturation ratio exceeds 0.8, even a right-turn percentage of only 10 percent could reduce delays substantially. This is not an unexpected result. Under pretimed control, the delays rise rapidly when the saturation ratio is greater than 0.8. At this level of the saturation ratio, the delays are about 25 sec or more per vehicle.

With RTOR

When RTOR is allowed from an auxiliary right-turn lane, right-turn delays may be further reduced. The extent of the reduction depends on the cross flow. When the cross flow is heavy and its saturation ratio approaches or exceeds 1.0, the gaps in this flow that are acceptable to the RTOR drivers would hardly exist. Consequently, RTOR would become virtually impossible. Under such circumstances RTOR cannot reduce right-turn delays. On the other hand, if the cross flow does not exist or is light, then the right-turn vehicles in a queue can execute multiple right-turns-on-red at a rate of about 1 vehicle per 4.7 sec (Figure 2). This could lead to a significant reduction in the delays.

The reductions in the right-turn delays attributable to RTOR also rest on the right-turn percentage. For example, with only 10 percent right turns in the approaching flow, RTOR has little influence on right-turn delays. With 40 percent right turns, then the availability of RTOR opportunities may reduce the delays substantially. This impact of RTOR at a 40 percent right-turn percentage is shown in Figure 6. Each of the curves in the figure represents the delays of a given right-turn flow under various cross-flow conditions. It can be seen from this figure that the average right-turn delays vary approximately in a linear manner with the saturation ratio of the cross flow. In this figure the average delays at a saturation ratio of 1.0 correspond closely to the average delays of right-turn vehicles

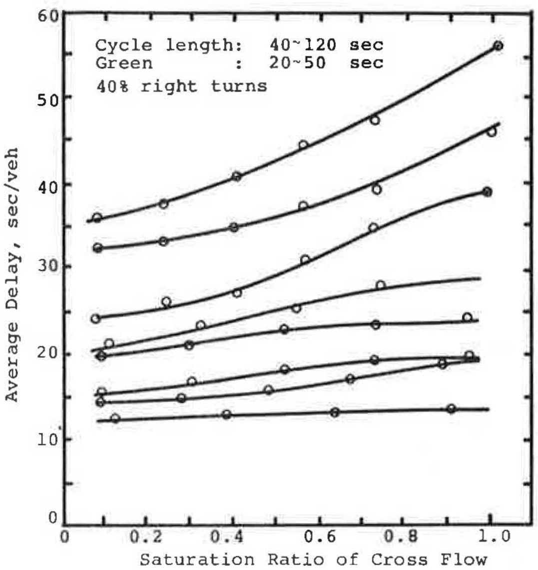


FIGURE 6 Effects of RTOR on right-turn delays.

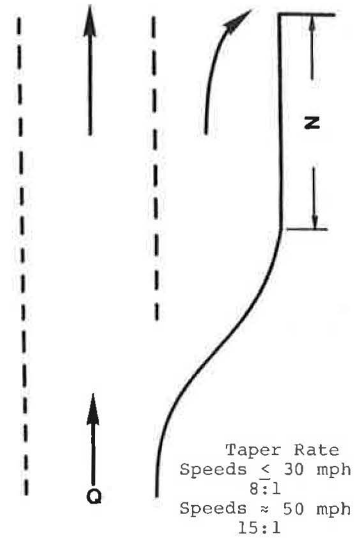


FIGURE 7 Schematic of auxiliary right-turn lane.

when RTOR is not allowed. These delays can also be conveniently estimated from Webster's formula (11) by using a saturation flow rate of 1,500 vph for a right-turn flow.

On the basis of the data in Figure 6, possible reductions in right-turn delays for an approaching flow with 40 percent right-turns are determined (Table 3). The data in this table indicate that RTOR has a negligible impact on delays if the average right-turn delays without RTOR are less than 15 sec per vehicle. Generally, RTOR is not likely to reduce right-turn delays significantly if the saturation ratio of the cross flow is greater than 0.6 and the delays without RTOR are less than 30 sec per vehicle.

TABLE 3 Reductions in Average Right-Turn Delays Due to RTOR, with 40 Percent Right Turns

Saturation Ratio[a]	Avg Delay Without RTOR (sec/vehicle)				
	0~15	15~20	20~30	30~45	>45
<0.2	0~2	0~5	5~8	8~15	>15
0.2~0.4	0	0~3	3~6	6~14	>14
0.4~0.6	0	0~2	2~4	4~12	>12
0.6~0.8	0	0~1	1~3	2~9	>9
0.8~1.0	0	0	0	0~5	>5

[a]Cross flow.

Approach lanes with rather high percentages of right turns are common. For example, out of the 16 lanes listed in Table 1, 12 had right-turn percentages in the range of 35 to 63 percent. The average for the 16 lanes was 39 percent. Note that none of these lanes has an auxiliary right-turn lane.

STORAGE REQUIREMENTS OF RIGHT-TURN LANES

Figure 7 shows an approach lane diverging into a straight-through lane and an auxiliary right-turn lane. The right-turn lane has a full-width section and a taper. As mentioned previously, AASHO recommended that neither the straight-through lane nor the right-turn lane be blocked. This requires that the queue lengths in both lanes and during any red phase be less than the storage length of the right-turn lane. The following variables should be considered when determining the minimum storage length of the right-turn lane needed to satisfy this requirement:

- Q = average flow rate in the approach lane that serves both straight-through and right-turn vehicles;
- R = length of red phase faced by the vehicles in the approach lane;
- λ = average number of arrivals during a red phase in the approach lane, equal to Q times R for a flow pattern with random arrivals;
- f = average right-turn flow rate as a proportion of Q $(0 < f < 1)$; and
- N = number of vehicles that can be stored in the full-width section of the right-turn lane during a red phase.

Assume that vehicles arrive randomly; therefore the probability of having x arrivals $(x = 0,1,2,\ldots)$ during a red phase can be approximated by the Poisson distribution (12):

$$P(x) = \lambda^x e^{-\lambda}/x! \tag{3}$$

where $P(x)$ is the probability of having x number of vehicles arriving during a red phase in the approach lane.

Given that there are x arrivals during a red phase, the probability of having y straight-through vehicles among these x arrivals is

$$F = \{x!/[(x-y)!y!]\}(1-f)^y f^{x-y} \tag{4}$$

where $x - y$ represents the number of right-turn vehicles.

For a right-turn lane with a full-width of 12 ft and a taper rate of 10:1, the taper would have a length of 120 ft. With this design feature, a blockage will rarely occur if the number of arrivals (x) is less than or equal to $N + 2$. By using this relationship as a basis for analysis, it can be assumed that a blockage will not occur if $x \leq N + 2$.

When $x > N + 2$, the straight-through vehicles will block the approach lane if $y > N + 2$. On the other hand, if $x - y > N + 2$, or if $y < x - N - 2$, the right-turn vehicles will do the blocking. For a given x, the probability of a blockage by the straight-through vehicles is

$$P_s = \sum_{y=N+3}^{x} \{x!/[(x - y)!y!]\} (1 - f)^y f^{x - y} \quad (5)$$

The probability of a blockage by the right-turn vehicles is

$$P_r = \sum_{y=0}^{x-N-3} \{x!/[(x - y)!y!]\} (1 - f)^y f^{x - y} \quad (6)$$

Equations 5 and 6 may share the same event. This event is associated with x = 2 (N + 3). In such a case the upper bound of y in Equation 6 is the same as the lower bound of y in Equation 5. This event has the following probability of occurring:

$$P_e = \sum_{y=0}^{N+3} \{(2N + 6)!/[(2N + 6 - y)!y!]\} \times (1 - f)^y f^{2N + 6 - y} \quad (7)$$

To determine the total probability of blockage, P_e should be accounted for only once. Therefore the total probability of blockage, taking into consideration all possible values of x greater than N + 2, is

$$P_t = \left[\sum_{x=N+3}^{\infty} P(x) (P_s + P_r) \right] - P_e \quad (8)$$

Because P_e is small in comparison with the sum of the probabilities of other events, it may be deleted from the equation.

To determine the storage requirement of a right-turn lane, different values of N in Equation 8 can be used to determine the probabilities of blockage. The smallest N that reduces the probability of blockage to an acceptable level is the minimum required storage capacity. Figure 8 shows the minimum storage capacities determined from Equation 8 for various combinations of λ and f to limit the probability of blockage to less than 0.1 percent. The same results can also be obtained through computer simulation.

Figure 8 shows that the required capacity for a given number of arrivals during a red phase is smallest when the right-turn vehicles account for 50 percent (i.e., f = 0.5) of the approaching flow. A larger capacity is needed when there is an uneven mix of straight-through and right-turn vehicles. The figure can be used easily to determine the minimum storage requirement of a right-turn lane.

For example, consider a case that involves the following signal control and traffic flow conditions: (a) total approach flow = 600 vph with 30 percent right turns, (b) cycle length = 60 sec, and (c) red phase faced by the right turns = 30 sec. Based on these data, the average combined number of straight-through and right-turn vehicles per red phase is λ = 600 x 30/3,600 = 5 vehicles. With λ = 5 and f = 0.3, the data in Figure 8 indicate that a minimum storage requirement of 8 vehicles is required for the full-width section of the right-turn lane.

The cost of providing an auxiliary right-turn lane varies with a number of factors. Therefore it was estimated that a right-turn lane with a 500-ft full-width section and a 20-ft taper would cost about $25,000. (Note that these data are from 1983

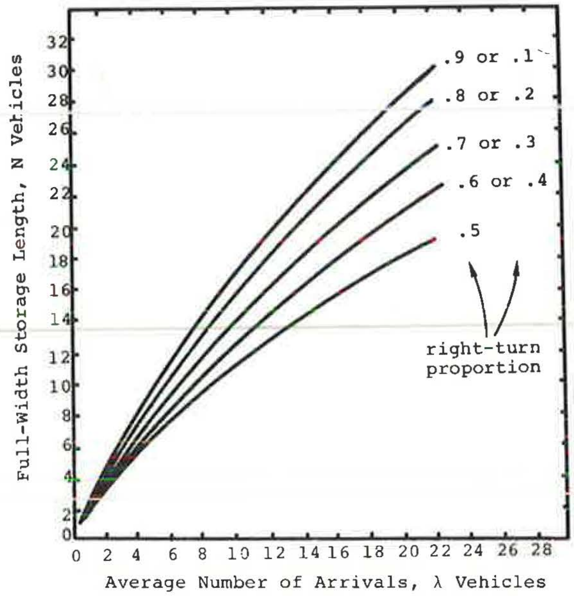


FIGURE 8 Minimum storage length requirements of auxiliary right-turn lanes.

correspondence with L. Raymond Powers, assistant engineer to the regional director of the Region 7 office of the New York State Department of Transportation in Watertown, New York.) This lane would have a 12-ft-wide flexible pavement with a 4-ft shoulder. The cost estimate allows a certain amount of earthwork. It reflects the probable cost of construction if the crews of a regional office of New York State Department of Transportation do the work.

As shown in Figure 8, the storage length of a right-turn lane can vary substantially with λ and f. For most urban intersections, the required lengths could be much shorter than the one mentioned previously.

It should be noted that Figure 8 is strictly valid only if the arrivals of vehicles are random. Nevertheless, it is adequate for most applications as long as the average number of arrivals (λ) is based on expected arrivals during a red phase. If the design flows are light (e.g., less than 200 vph per lane during red phases), the arrivals could have larger variations than those in a random arrival pattern. As a result, storage lengths longer than those shown in Figure 8 may be needed. This is not a serious problem because an insufficient storage length under a light flow condition would not have a significant adverse impact on the traffic operation. Besides, the storage lengths for such a case may be lengthened slightly from the lengths shown in Figure 8. If the design flows are heavy (e.g., more than 700 vph per lane during red phases), the arrivals could have smaller variations. The use of Figure 8 in such events could lead to conservative storage lengths.

CONCLUSIONS

The use of auxiliary right-turn lanes to accommodate RTOR can complement improved signal controls to increase the efficiency of traffic operations in a street network. Current practices in intersection traffic operations often require straight-through vehicles to share a lane with right-turn vehicles. Field data indicate that in such a lane it is not unusual for right turns to constitute more than 35 percent of the traffic. If the average delay of vehicles in a mixed straight-through and right-turn flow exceeds approximately 25 sec per vehicle, the provision of an auxiliary right-turn lane can substantially reduce the delays even if the right-turn percentage is only 10 percent.

To facilitate RTOR, the availability of auxiliary right-turn lanes is indispensable. Without an auxiliary right-turn lane, it has been observed that a limited shoulder area can allow more than 6 times as many vehicles to turn on red as a regular traffic lane. This signifies the desirability of providing auxiliary right-turn lanes to accommodate RTOR.

The critical gap of RTOR drivers was found to be approximately 8.4 sec. This is about 70 percent longer than the critical gap of opposed left-turn drivers. When a long gap is available in the cross flow, multiple RTOR requires an average of about 4.7 sec per vehicle to complete. As a result, a red phase is only 60 percent as useful as a green phase of the same length even when the cross flow does not exist.

RTOR from an auxiliary right-turn lane may not reduce right-turn delays significantly. When the right-turn delays without RTOR are less than 15 sec per vehicle, allowing RTOR would have negligible effects on the delays. RTOR could effectively reduce right-turn delays if the saturation ratio of the cross flow is less than 0.6 and the right-turn delays without RTOR exceed 30 sec per vehicle.

The storage length of an auxiliary right-turn lane should be long enough to prevent a blockage of traffic lanes during a red phase. The minimum storage requirements depend primarily on the flow rate and the right-turn percentage of an approach flow. A design chart that relates the minimum storage requirements to these influencing factors is presented in this paper.

ACKNOWLEDGMENTS

The author wishes to express his appreciation to Martin C. Percy, Frederick J. Wagner, and Robert Hopkins for their assistance in the data collection. Thanks are also extended to L. Raymond Powers of the New York State Department of Transportation for providing the cost data on auxiliary right-turn lanes.

REFERENCES

1. E.B. Lieberman and S. Cohen. New Technique for Evaluating Urban Traffic Energy Consumption and Emissions. In Transportation Research Record 599, TRB, National Research Council, Washington, D.C., 1976, pp. 41-45.
2. M.-F. Chang, L. Evans, R. Herman, and P. Wasielewski. Fuel Consumption and Right Turn on Red: Comparison Between Simple Model Results and Computer Simulation. Letter to the Editor, Transportation Science, Vol. 11, No. 1, 1977, pp. 92-94.
3. J.E. Clark, S. Maghsoodloo, and D.B. Brow. Public Good Relative to Right-Turn-on-Red in South Carolina and Alabama. In Transportation Research Record 926, TRB, National Research Council, Washington, D.C., 1983, pp. 24-31.
4. M.S. Mamlouk. Right-Turn-on-Red: Utilization and Impact. Final Report, Project C-36-1700. Joint Highway Research Project, Purdue University, West Lafayette, Ind., 1976.
5. R.L. May. RTOR: Warrants and Benefits. Research Report 14. Joint Highway Research Project, Purdue University, West Lafayette, Ind., 1974.
6. H.W. McGee, W.A. Stimpson, J. Cohen, G.F. King, and R.F. Morris. Right Turn on Red, Volume I. Final Technical Report FHWA-RD-89. FHWA, U.S. Department of Transportation, 1976.
7. A Policy on Design of Urban Highways and Arterial Streets. AASHO, Washington, D.C., 1973, pp. 688-689.
8. J. Behnam. Gap Acceptance as a Criterion for Left-Turn Phasing. Traffic Engineering, Vol. 42, No. 9, June 1972, pp. 40-42.
9. O.K. Dart, Jr. Left-Turn Characteristics at Signalized Intersections on Four-Lane Arterial Streets. In Highway Research Record 230, HRB, National Research Council, Washington, D.C., 1968, pp. 45-59.
10. 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.
11. F.V. Webster. Traffic Signal Settings. Road Research Technical Paper 39. Great Britain Road Research Laboratory, London, England, 1958.
12. D.L. Gerlough and M.J. Huber. TRB Special Report 165: Traffic Flow Theory. TRB, National Research Council, Washington, D.C., 1975, p. 17.

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

Determination of Motorist Violations and Pedestrian-Related Countermeasures Related to Right-Turn-on-Red

CHARLES V. ZEGEER and MICHAEL J. CYNECKI

ABSTRACT

The adoption of the permissive right-turn-on-red (RTOR) rule in the United States (except for New York City) has resulted in possible problems with respect to motorists failing to make a full stop before turning right on red. Also, the rate of motorist violations to the NO TURN ON RED sign was also raised as a related problem. The purpose of this study was to determine these violation rates relative to RTOR and to determine the resulting pedestrian-related conflicts associated with RTOR maneuvers. Based on the collection of observational data for more than 67,000 drivers at 110 intersections in Washington, D.C.; Detroit, Michigan; and Dallas and Austin, Texas, 3.7 percent of all right-turning motorists at RTOR-prohibited intersections violate the RTOR prohibition signs. However, of those motorists given an opportunity to commit an RTOR violation, about 21 percent violate the NO TURN ON RED sign. Although 23.4 percent of RTOR violations result in a conflict with another vehicle or pedestrian, only about 1 out of every 100 total right-turn vehicles is involved in an RTOR conflict. In terms of stopping characteristics at RTOR-allowed sites, 56.9 percent of motorists fail to make a full stop before turning right on red. An analysis of specific data-collection sites resulted in a list of locational factors associated with high and low violation rates. From this analysis a list of 30 candidate countermeasures was developed for possible use relative to RTOR.

The recent adoption of the Western Rule in the United States relative to right-turn-on-red (RTOR), except for New York City, has resulted in the right of motorists to turn right on a red signal (except when otherwise signed) after stopping and yielding to pedestrians and motorists. However, two of the reported problems of the generally permissive RTOR rule involve motorists:

1. Turning right on red at RTOR-prohibited locations (i.e., NO TURN ON RED signs exist), and
2. Turning right on red (where permitted) without stopping.

It has been speculated that one of the causes of violations of RTOR prohibitions is the carry-over effect to motorists because of the current permissive RTOR rule that causes them to expect to be able to turn right on red at all intersections. One confounding problem is that the NO TURN ON RED (NTOR) sign is not always placed in the same position, and it may not be noticeable to drivers even when the sign is placed in accordance with Manual on Uniform Traffic Control Devices (MUTCD) standards (1). Other problems involve the lack of police enforcement of RTOR prohibition in many areas. The current MUTCD warrants for an NTOR sign have led to the high use of RTOR prohibitions in some cities and little or no use in other cities. Many believe that RTOR is not hazardous, and therefore prohibitions are rarely if ever needed. Others view RTOR as a detriment to safety in that it should never have been implemented.

The other compliance problem with RTOR relates to RTOR vehicles that fail to come to a full stop before turning right on red where RTOR is allowed. Previous studies have indicated that between 3 and 65 percent of vehicles commit such RTOR violations (2,3). However, only about 1 to 3 percent of RTOR violations (i.e., failing to stop) resulted in an unsafe act or hazardous situation (3).

With evidence of these two types of RTOR violations, a need exists to determine the current status of motorist compliance with RTOR prohibition. Therefore, the purpose of this study was to

1. Conduct observational studies at signalized intersections in several cities to determine current motorist compliance with RTOR prohibition (NTOR signs) and the requirement to make a full stop before turning right on red (where RTOR is permitted);
2. Collect traffic, geometric, and other physical site characteristics and determine what site factors are associated with high and low rates of RTOR violations; and
3. Develop a list of countermeasures for increasing compliance or reducing hazards or both related to RTOR.

MOTORIST COMPLIANCE WITH RTOR LAWS

One of the objections to the generally permissive RTOR regulation is that motorists frequently do not stop before turning on red. Such concerns have recently been expressed in several studies (2-5). An assessment of motorist compliance with stopping is presented in the following section, followed by a discussion of motorist violation of turning on red where the maneuver is prohibited.

Compliance Where RTOR is Permitted

The generally permissive RTOR rule requires that motorists must come to a full stop and yield to pedestrians and other traffic in the intersection before turning on red. There have been several examinations of motorist compliance and violations to the RTOR law. A 1983 study (6) found that overall, 40 percent of the drivers who turned on red failed to come to a stop before turning. Violation rates per site ranged from 38 to 71 percent of RTOR vehicles. Under the sign-permissive rule in Virginia, Parker et al. (7) found that 9 percent of the RTOR motorists at 15 approaches did not come to a full stop before turning. A study conducted at 11 sites in Providence, Rhode Island, found that 65 percent of the motorists did not stop (8). At 12 locations in Springfield, Massachusetts, only 28 percent of the RTOR motorists did not come to a full stop (8). The low violation rate in Springfield was attributed to the newness of the RTOR maneuver and the sign reminding motorists to stop. Baumgaertner (3) collected compliance data at 13 approaches in Maryland and also found that the noncompliance rate under the sign-permissive rule was 64.4 percent, which compares closely with the Providence data.

RTOR violation data were collected for generally permissive RTOR in two studies in which the general rule had only been adopted for 1 year (2,6). At seven approaches in North Carolina, Parker et al. (7) found that 2.0 percent of the RTOR motorists did not stop. However, after generally permissive legislation was enacted in Virginia, Parker (2) found that 11.5 percent of the RTOR motorists violated the law. It is important to note that the violation rate varied considerably with 48 percent of the violations reported at two approaches.

A high violation rate creates a law enforcement problem and may lead to a serious safety problem. In their studies, Baumgaertner (3) and Parker (2) also recorded the number of unsafe turns where the RTOR motorists did not stop or yield to other traffic in the immediate vicinity of the intersection. In both studies less than 2 percent of the motorists made an unsafe turn. Additional studies of motorist compliance are needed periodically to examine trends over time and to identify unsafe approaches so that appropriate countermeasures can be applied.

The magnitude of the RTOR violation problem can be put into perspective by comparing it with motorist compliance at stop sign locations. In a Chicago study, 53 to 76 percent of all drivers failed to come to a complete stop at stop signs. However, only 5 to 10 percent of all vehicles traveling in excess of 5 mph (8 km/h) violated the stop sign (9). A 1976 study by Baubien (10) was conducted in Troy, Michigan, to determine whether stop signs were effective for speed control in residential areas. At the three locations full stops ranged from 6 to 51 percent of vehicles, rolling stops ranged from 34 to 54 percent, and no-stops ranged from 15 to 47 percent (10). Based on these data, the violation rate involving stop signs appears to be considerably higher than the RTOR noncompliance rate.

A 1978 study observed motorist obedience to the stop signs in Barton, Springfield, and Providence. The percentage of vehicle violations (not stopping) ranged from 31 to 39 percent. Of those vehicles not forced to stop by cross-street traffic, the percentage of violations (nonstopping vehicles) ranged from 35.2 to 71.2 percent (8).

Violations Where RTOR is Prohibited

Another major concern is whether motorists are violating the law by turning right on red at locations where the maneuver is prohibited. There is evidence that violations do occur. The most recent study was conducted in New Jersey in 1983, and it was noted that 6 percent of right-turn vehicles turned on red (at five intersections) where RTOR was prohibited (6).

Benke and Ries (11) collected violation data at 11 sites where RTOR maneuvers were prohibited under sign-permissive and generally permissive rules and found that the violation rates were 1.23 and 9.56 percent, respectively (i.e., 1.23 percent of the motorists made an illegal RTOR maneuver). The authors attributed the high violation rate, which occurred at 4 of the 11 sites, to poor visibility of the sign resulting from poor sign placement and a busy signing environment at one location. In Indiana Mamlouk (12) found that 1.4 percent of the motorists made an illegal RTOR maneuver under the sign-permissive rule. It was also reported that the violation rate varied considerably, with one site having an 18 percent violation rate. At that location sign placement made it difficult for motorists to see the traffic control device.

METHODOLOGY

Data were collected to investigate two problems associated with RTOR: (a) to determine if RTOR prohibitions are being obeyed, and (b) to determine if motorists are coming to a complete stop before making an RTOR maneuver where RTOR is permitted. Each of these problems required separate data-collection plans and procedures, as discussed in the following sections.

Data-Collection Plan for Violations of RTOR-Prohibited Locations

Selection of Cities for Data Collection

One of the factors that could have a major impact on RTOR compliance is the recent history of RTOR in the area, because this could influence the level of motorist knowledge and understanding of RTOR and RTOR prohibition. For example, motorists in cities that have had the Western Rule for many years (e.g., Los Angeles) may respond differently to RTOR prohibition than motorists in eastern cities that have used the Eastern Rule until recently (e.g., Washington, D.C.). Other factors such as level of police enforcement of RTOR, area characteristics, and local driver characteristics may also affect the level of compliance and vary from city to city or state to state, although such factors are difficult or impossible to quantify.

To allow for collecting data for a variety of conditions, three U.S. metropolitan areas were selected:

1. One city in the western United States that has had the Western Rule (RTOR permissive law) in effect for many years,
2. One city in the eastern United States that has only recently adopted the Western Rule (within 4 or 5 years), and
3. One city in a neutral part of the country such as the Midwest.

After discussions with the FHWA and numerous cities, it was decided to use Washington, D.C., to represent the city that until recently had the Eastern Rule. The cities of Dallas and Austin, Texas, were selected to represent cities with the Western Rule, and Detroit, Michigan, was selected from the

Midwest. Washington, D.C., currently prohibits RTOR (for either part of the day or all day) at approximately 70 percent of its intersections. RTOR is prohibited at only a small percentage of intersections in Dallas and Austin, whereas RTOR prohibition is used at an estimated 10 to 20 percent of signalized intersections in the Detroit area.

Selection of Data-Collection Sites

Sites were selected to provide a variety of geometric, volume, and other conditions throughout the city. One of the site selection criteria was moderate to high levels of pedestrian volume. However, some sites with low pedestrian volumes were selected that exhibited unusual geometrics. Also, intersections that have two or more approaches that prohibit RTOR were selected in many instances to facilitate data collection.

To select the sites and approaches, a list of sites with RTOR prohibition was obtained from each city. The sites were field reviewed by the project engineers before data collection. During this review basic site information was obtained and observation points and data-collection time periods were selected. Violation data were collected for a total of 110 approaches to provide a variety of site characteristics.

Development of Data-Collection Forms and Procedures

Data-collection forms and procedures were developed to assist observers in obtaining accurate and consistent data. Two basic types of data were collected: site data and violation data. Site data collected included all traffic control devices (signs, signals, and pavement markings), intersection geometrics, posted speed limits, sight distance for the right-turn vehicle, and pertinent signal data.

The reverse side of the form was used for the condition diagram, and observers were instructed to draw a detailed site diagram with street widths, location of pavement markings, signs and signals, special turn lanes, intersection geometry, type of development on each corner, location of on-street parking (if any), and other physical features. Observation data were collected in 10-min intervals on form 1 and included the following items:

1. Start time and end time of the data-collection period (military time).
2. Approach (northbound, eastbound, and so forth).
3. The number of right-turn-on-green (RTOG) vehicles. RTOG vehicles were categorized into arrive on green, arrive on red (RTOR opportunity), and arrive on red (no RTOR opportunity).
4. RTOR maneuvers, which were categorized into no conflict, conflict with traffic, and conflict with pedestrians. Pedestrian conflicts were recorded based on whether they occurred at the near or far crosswalk and the type of conflict: (a) vehicle hesitation (VH)--vehicle slows or stops to avoid hitting a pedestrian while executing an RTOR maneuver; (b) vehicle swerve (VS)--vehicle swerves to avoid hitting a crossing pedestrian; (c) pedestrian hesitation (PH)--pedestrian slows, stops, or reverses his direction of travel to avoid a collision; (d) pedestrian run (PR)--pedestrian increases his speed or runs to avoid a collision; and (e) interaction (I)--neither the vehicle nor the pedestrian reacts but the pedestrian is in a moving lane and is within 20 ft (6 m) downstream of an RTOR vehicle.
5. Pedestrian volume, where the total number of crossing pedestrians is recorded separately for the near and far crosswalks, regardless of their direction of travel or compliance with the pedestrian or traffic signal.

When two or more conflict types occurred during a single event (i.e., a vehicle hesitates and a pedestrian runs during the same RTOR event), only the most severe conflict was recorded. Only one conflict was recorded per RTOR vehicle, regardless of the number of pedestrians involved in the conflict.

A minimum of 4 hr of data was collected on each approach. Eight or more hours of data were collected on several approaches to test for data repeatability.

Data-Collection Plan for Violation Data at RTOR-Permitted Approaches

This portion of the study involved collecting violation data at RTOR-permitted sites to determine whether vehicles were making a complete stop before their RTOR maneuver. These data were later compared with stopping characteristic data for right-turn motorists at stop sign locations. The data were collected at sites within Washington, D.C.; Dallas/Austin, Texas; and Detroit, Michigan, as discussed earlier.

Selection of Data-Collection Sites

Sites selected included signalized intersections with at least two approaches that permit RTOR or intersections with at least two approaches controlled by stop signs. Initial site selection was made by selecting a list of potential test sites. Final site selection was made by reviewing candidate sites with high right-turn volume, high RTOR volume (signalized locations), and moderate to high pedestrian volumes. The sites selected were in the vicinity of the RTOR-prohibited locations used for collection of violation data relative to prohibition signs. Data were collected for 29 total approaches of signalized intersections and 28 stop sign approaches.

Development of Data-Collection Forms and Procedures

Data collected included site information and stopping characteristics (observation data). Site data were also collected as described earlier. Observation data were collected on the RTOR and stop sign stopping characteristics data form. A total of 4 hr of data were collected on each approach, or a total of 8 hr at each intersection. Data collection was alternated between two approaches with 30 min of data collected on an approach (summarized and recorded in 10-min intervals). In this manner, data were sampled from both approaches throughout the day.

Data collected on the RTOR and stop sign stopping characteristics data form included the following:

1. Intersection name, city, location, and so forth.
2. Intersection control, such as traffic signal or stop sign.
3. Time period data collection began and ended (military time).
4. Approach (northbound, eastbound, and so forth).
5. RTOG--the number of vehicles that turn right on green signal indications (for signalized approaches only).

6. RTOR vehicles--the type of stop for RTOR or stop sign right-turn vehicles, which are defined as (a) no stop--the vehicle slows only to negotiate the right turn and does not make any effort to stop; (b) rolling stop--the right-turn vehicle slows more than the no-stop condition but at no time do the wheels come to a complete stop in the vicinity of the stop bar or crosswalk; (c) full stop-voluntary--the vehicle comes to a complete stop in the vicinity of the stop bar or crosswalk but is not forced to stop by pedestrians in the crosswalk or by cross-street traffic; and (d) full stop-forced--the vehicle comes to a complete stop in the vicinity of the stop bar or crosswalk and does so because of the existence of pedestrian crosswalk activity or through traffic. (Note that this does not necessarily mean the vehicles would not have voluntarily stopped if no pedestrian or cross-traffic were present.)

7. Pedestrian volume--crossing pedestrian traffic on the near or far side crosswalk.

8. Opposing traffic--the cross traffic potentially conflicting with RTOR or right-turns at stop signs. For an approach that intersects a two-way street, only the direction of cross traffic that conflicts with the right-turn maneuver would be counted.

RESULTS

Status of Violations to RTOR Prohibition Signs

Violation data were collected at a total of 110 intersection approaches relative to vehicles illegally turning right on red. The violation rate for a group of sites may be expressed in several different ways:

1. Overall RTOR violation rate is the overall percentage of right-turn vehicles that turn right on red (i.e., total number of RTOR events at a group of sites divided by the total right-turn volume). This was a common way of expressing violations in past studies.

2. Mean RTOR violation rate is the average percentage of right-turn vehicles that turn right on red (i.e., the mean percent violations of a sample of intersection approaches). This can only be computed for a sample of two or more sites.

3. Overall RTOR violation rates per opportunity is the percentage of vehicles turning right on red of those vehicles that have an opportunity to do so. In the first two definitions (1 and 2), all right-turning vehicles are included in the denominator, regardless of whether they arrive on red, arrive on green, or had an opportunity to make an RTOR (i.e., they were the second or third car stopped in the right-turn lane, or a lack of gaps in cross-street traffic prevented them from turning right on red). This definition only includes those vehicles stopped first in line at the red light that have an adequate gap and an opportunity to turn right on red. It is really a measure of the percentage of motorists who would violate the RTOR prohibition if given the chance. This definition will result in a higher percent violation rate than the previous two definitions.

4. Mean RTOR violation rate per opportunity is the same as the previous definition, except a mean of the violation rates of the sites is used.

To illustrate the three definitions of violation rate, consider hypothetical data on three intersection approaches, A, B, and C (1 hr of data per approach) when each has NTOR signs (Table 1). From the sample data in Table 1, the overall RTOR violation rate for the three approaches is the total RTOR (18) divided by the total right turns (135), or 13.3 percent. The mean RTOR violation rate for the three approaches is the average of 6.0 percent (Approach A), 11.1 percent (Approach B), and 25.0 percent (Approach C), or 14.0 percent. This differs slightly from the 13.3 percent overall RTOR violation rate.

To compute the overall and mean RTOR violation rate per opportunity only the RTOR opportunities are used in the denominator. Thus, in the sample data in Table 1, the overall RTOR violation rate per opportunity for the three approaches is the total number of violations (18) divided by the total opportunities (60), or 30.0 percent. The mean RTOR violation rate per opportunity is computed as the average violation rate of Approach A (30.0 percent), Approach B (25.0 percent), and Approach C (33.3 percent), or 29.4 percent, which differs slightly from the 30.0 overall rate.

The actual violation rates are given in Table 2 for each of the three cities and for the overall data base. Of the 110 intersection approaches, 59

TABLE 1 Hypothetical Data on Three Intersection Approaches

Approach	Total Right Turns	RTOR Violations	RTOR Opportunities	Vehicles Turning Right on Red (%)	Vehicles Turning Right on Red That Had An Opportunity (%)
A	50	3	10	6.0	30.0
B	45	5	20	11.1	25.0
C	40	10	30	25.0	33.3
Total	135	18	60		

TABLE 2 Summary of RTOR Violations at RTOR-Prohibited Sites

City	Total Approaches	Right Turns	Total RTOR Violations	Violation Rate (%)		Total RTOR Opportunities	Violation Rate per Opportunity (%)	
				Overall	Mean		Overall	Mean
Detroit	59	33,400	1,119	3.4	4.7	5,904	19.0	22.0
Washington, D.C.	27	22,742	888	3.9	4.6	4,122	21.5	19.4
Dallas/Austin	24	11,205	493	4.4	6.9	2,288	21.5	24.6
Total	110	67,347	2,500	3.7	5.1	12,314	20.3	21.9

TABLE 3 Summary of Violations and Conflicts at RTOR-Prohibited Sites

City	Total No. of Violations	RTOR Violations Resulting in Conflicts: Total Conflicts No.	Total Conflicts Percent	Conflicts with Traffic No.	Conflicts with Traffic Percent	Total Conflicts with Pedestrians No.	Total Conflicts with Pedestrians Percent	Pedestrian Conflicts: Near Crosswalk Only No.	Near Crosswalk Only Percent	Far Crosswalk Only No.	Far Crosswalk Only Percent
Detroit	1,119	246	22.0	79	7.1	167	14.9	61	5.5	106	9.5
Washington, D.C.	888	199	22.4	28	3.2	171	19.3	44	5.0	127	14.3
Dallas/Austin	493	140	28.4	80	16.2	60	12.2	34	6.9	26	5.3
Total	2,500	585	23.4	187	7.5	398	15.9	139	5.6	259	10.4

were from Detroit, 27 from Washington, D.C., and 24 from the Dallas/Austin area. A total of 2,500 violations were observed for the 67,347 total turning vehicles, or 3.7 percent overall. The overall violation rates ranged between 3.4 percent (Detroit) and 4.4 percent (Dallas/Austin). The mean violation rate was 5.1 for all sites and ranged from 4.6 percent (Washington, D.C.) to 6.9 percent (Dallas/Austin). These numbers compare closely with the 6 percent overall violation rate found by Davis and Mullowney (6) in New Jersey at 11 sites in a 1983 study.

Other information in Table 2 relates to RTOR violation rates per opportunity. For example, of the 67,347 right turns at the 110 sites, only 12,314 (18.3 percent) had an opportunity to turn right on red. This is because many arrived and turned right on green or were not the lead vehicle stopped in the right-turn lane (could not physically make the turn on red). In a few cases no opportunity existed for a RTOR violation because of high pedestrian or cross-street traffic.

The overall RTOR violation rate per opportunity was 20.3 percent. The rate was consistent among the cities, ranging from 19.0 percent (Detroit) to 21.5 percent (Washington and Dallas/Austin). This indicates that about 1 out of every 5 motorists turns right on red when given the opportunity when it is prohibited.

One additional analysis was also conducted of the percentage of overall RTOR violations that resulted in a conflict, as summarized in Table 3. Of the 2,500 total RTOR violations at the 110 approaches, 585 (23.4 percent) resulted in some type of conflict. Of the 2,500 violations, 187 (7.5 percent) involved cross traffic, 139 (5.6 percent) involved pedestrians in the near crosswalk, and 259 (10.4 percent) involved pedestrians in the far crosswalk.

In terms of individual cities, RTOR violations in Dallas/Austin resulted in a conflict 28.4 percent of the time compared with approximately 22 percent in the other two cities. In particular, 16.2 percent of RTOR violations in Dallas/Austin resulted in a cross-traffic conflict, compared with 3.2 percent and 7.1 percent in Washington, D.C., and Detroit, respectively. However, pedestrian-related conflicts ranged from 19.3 percent of RTOR violations in Washington, D.C., compared with 14.9 percent (Detroit) and 12.2 percent (Dallas/Austin), probably because of the higher densities of pedestrians at the Washington sites.

These pedestrian conflicts occurred most frequently on the near crosswalk in Dallas/Austin (6.9 percent on the near crosswalk to 5.3 percent on the far crosswalk). However, the far crosswalks experienced more pedestrian conflicts than the near crosswalks at the sites in Washington (14.3 to 5.0 percent) and Detroit (9.5 to 5.1 percent). RTOR violations with pedestrians in the far crosswalk could be largely the result of pedestrian violations, because during a red phase pedestrians in the near crosswalk would normally have the WALK interval.

It should be remembered from the previous discussion that although 23.4 percent of all RTOR violations resulted in conflicts, only 3.7 percent of all right-turning vehicles committed an RTOR violation. Thus only 0.234 x 0.037 = 0.9 percent (less than 1 in 100) of the right-turn vehicles was involved in any kind of an RTOR-related conflict (585 RTOR-related conflicts for 67,347 total right-turning vehicles). Further, RTOR-pedestrian conflicts resulted from only 398 of 67,347 right-turning vehicles (0.59 percent), or about 6 out of every 1,000 right-turning vehicles. It should also be remembered that a majority of the sample sites were in areas with moderate to high pedestrian volumes, so these percentages of pedestrian conflicts are likely higher than would be expected for the overall sample of intersections in a city.

As discussed earlier, details were also recorded for the specific types of pedestrian conflicts resulting from each RTOR violation, as summarized in Table 4. Of the 398 resulting pedestrian conflicts, the most prevalent types were pedestrian-vehicle interactions (36.5 percent), pedestrian hesitations

TABLE 4 Summary of Types of Pedestrian Conflicts Resulting from Violations of RTOR Prohibitions

Type of Pedestrian Conflict	Conflicts: Near Crosswalk No.	Near Crosswalk Percent	Far Crosswalk No.	Far Crosswalk Percent	Total No.	Total Percent
Vehicle hesitation	27	19.4	81	31.3	108	27.1
Vehicle swerve	2	1.5	4	1.5	6	1.5
Pedestrian hesitation	48	34.5	75	29.0	123	30.9
Pedestrian run	4	2.9	12	4.6	16	4.0
Pedestrian/vehicle interaction	58	41.7	87	33.6	145	36.5
Total	139	100.0	259	100.0	398	100.0

TABLE 5 Comparison of Pedestrian Conflicts Occurring with RTOR and RTOG

City	Total Right Turns	RTOR	RTOG	RTOR (%)	RTOR with Conflict: Cross Traffic No.	Cross Traffic Percent	Pedestrians at Near Crosswalk No.	Pedestrians at Near Crosswalk Percent	Pedestrians at Far Crosswalk No.	Pedestrians at Far Crosswalk Percent	RTOG with Conflict: Pedestrians at Near Crosswalk No.	Pedestrians at Near Crosswalk Percent	Pedestrians at Far Crosswalk No.	Pedestrians at Far Crosswalk Percent
Detroit	20,867	761	20,106	3.6	49	6.4	39	5.1	60	7.9	149	0.7	3,547	17.6
Washington, D.C.	9,000	334	8.666	3.7	5	1.5	17	5.1	57	17.1	87	1.0	2,628	30.3
Dallas/Austin	8,095	393	7,702	4.9	72	18.3	20	5.1	19	4.8	35	0.5	690	9.0
Total	37,962	1,488	36,474	3.9	126	8.5	76	5.1	136	9.1	271	0.7	6,865	18.8

(30.9 percent), and vehicle hesitations (27.1 percent). Only 16 pedestrian runs and 6 vehicle swerves were observed during the 573 hr of data collection. Vehicle hesitations were more prevalent in the far crosswalk than the near crosswalk (31.3 percent to 19.4 percent) and pedestrian-vehicle interactions were more common on the near crosswalk than on the far crosswalk (41.7 to 33.6 percent).

A comparison was also made between RTOR-related conflicts and RTOG conflicts for a sample of the data sites, as summarized in Table 5. The sample includes 37,962 right-turn vehicles, of which 1,488 (3.9 percent) illegally turned right on red and 96.1 percent turned right on green. In terms of pedestrians, 14.2 percent of RTOR maneuvers resulted in a pedestrian conflict compared with 19.5 percent of RTOG maneuvers that resulted in pedestrian conflicts. However, an additional 126 RTOR maneuvers (8.5 percent) resulted in cross-traffic conflicts. Thus a total of 22.7 percent (14.2 + 8.5) of illegal RTOR maneuvers resulted in a conflict, compared with 19.5 percent of RTOG conflicts. Thus, although illegal RTOR maneuvers result in a slightly higher rate of total conflicts than RTOG (22.7 to 19.5 percent), fewer pedestrian conflicts occurred with illegal RTOR maneuvers than with RTOG (14.2 percent compared with 19.5 percent). It should be mentioned that pedestrians may legally cross the street in the near crosswalk with RTOR and the far crosswalk with RTOG.

Status of Violations to the Stopping Requirement at RTOR-Permitted Sites

Data were collected at 29 RTOR-allowed approaches in the three cities relative to the frequency of vehicles making a full stop, rolling stop, or no stop when turning right on red, as summarized in Table 6. In addition, stopping data were also collected at 28 stop sign locations for comparison purposes. A total of 4 hr of data were collected per approach, for a total of approximately 228 hr of data. Conflict data were not collected relative to stopping characteristics data.

For the 29 signalized approaches (with RTOR allowed), 26.2 percent of right-turn vehicles turned right on red overall, with a small variation between cities (from 24.2 percent in Dallas/Austin to 29.3 percent in Washington, D.C.). Of all the vehicles turning right on red at the 29 approaches, 14.8 percent were recorded as no-stops (turned as if a green light existed), 42.1 percent made rolling stops, and 43.1 percent made full stops. Thus 56.9 percent (42.1 + 14.8 percent) of motorists violated the RTOR law by not making a full stop before turning right on a red signal. Of the 43.1 percent full stops, 36.0 percent were forced to stop (i.e., by oncoming traffic or pedestrians) and 7.1 percent were voluntary stops.

An analysis by city revealed that total violations (no-stops plus rolling stops) were the highest in Washington, D.C. (with 61.4 percent of vehicles not fully stopping) and Detroit (59.1 percent of vehicles not fully stopping), and lowest in Dallas/Austin (50.3 percent of vehicles not fully stopping).

The percentage of right-turning vehicles stopping at RTOR-allowed sites was compared with those at stop sign locations, because motorists under both situations are required to make a full stop and then turn right after yielding to pedestrians and cross-street traffic. Thus the relative magnitude of nonstopping motorists at RTOR-allowed locations could be discussed in terms of another type of traffic control. Such comparisons of compliance between RTOR-allowed sites and stop sign locations have been made in several previous RTOR studies.

The overall violation rate (i.e., motorists not fully stopping) of right-turn vehicles was found to be 68.2 percent at stop sign locations compared with 56.9 percent at the RTOR-permitted sites, a difference of 11.3 percent. Rolling stops were higher at the stop sign locations (57.3 percent) compared with RTOR-allowed locations (42.0 percent). However, the percentage of no-stops was 14.8 percent at the RTOR-permitted locations, compared with 10.9 percent at the stop sign locations.

The overall percentage of voluntary stops was approximately 7 percent at both the RTOR-allowed

TABLE 6 Summary of Data Collected at RTOR-Permitted and Stop Sign Approaches

Approach	Right Turns per Hour	RTOR per Hour	RTOR (%)	Stopping Violations per Hour	Stopping Violations (%): Total Violations	Rolling Stop	No Stop	Full Stops (%): Total	Voluntary	Forced	No. of Approaches
RTOR-allowed approaches (total)	67.3	16.3	26.2	9.2	56.9	42.0	14.8	43.1	7.2	36.0	29
Detroit	64.1	15.1	25.0	9.3	59.1	46.5	12.6	40.9	8.5	32.4	9
Washington, D.C.	69.3	19.5	29.3	11.7	61.4	41.7	19.7	38.6	4.6	34.1	10
Dallas/Austin	68.0	14.1	24.2	6.7	50.3	38.4	11.9	49.7	8.7	41.0	10
Stop sign approaches (total)	38.3		NA	27.1	68.2	57.3	10.9	31.8	7.1	24.7	28
Detroit	59.3		NA	43.5	67.3	56.5	10.8	32.7	6.0	26.7	10
Washington, D.C.	35.5		NA	22.5	63.0	49.5	13.5	37.0	10.1	26.9	8
Dallas/Austin	19.5		NA	14.3	73.3	64.3	8.9	26.7	5.9	20.8	10

Note: NA = not applicable.

sites and the stop sign locations. However, 36 percent of the RTOR motorists were forced to stop at the RTOR-allowed locations compared with 24.7 percent at the stop sign locations, a difference of 11.3 percent. Note that a difference of 11.3 percent was also found between RTOR-allowed and stop sign approaches in terms of overall violations. This indicates that the slightly higher percentage of vehicles stopping at the RTOR locations (43.1 percent) compared with the stop sign locations (31.8 percent) could be largely the result of more opportunities for a rolling or no stop at the stop sign locations. Thus it appears that there is little difference in driving behavior in terms of stopping compliance between the RTOR-permitted locations and the stop sign locations.

The overall 56.9 percentage of vehicles not fully stopping (before turning right on red) is higher than the 40 percent found by Davis and Mullowney (6) in a 1983 study of intersections in New Jersey. Part of the differences could be slight variations in the definitions of a rolling or full stop, differences in site characteristics, or differences in motorist behavior at the New Jersey sites. However, a 1978 study of 11 sites in Providence, Rhode Island, and 12 locations in Springfield, Massachusetts, found that 65 and 28 percent of the motorists, respectively, did not stop before turning right on red. The high compliance rate in Springfield was attributed to the newness of the RTOR maneuver and the sign reminding them to stop (8). In a 1981 study, Baumgaertner (3) found that 64.4 percent of drivers failed to stop in Maryland before turning right on red. Thus other recent studies have found rates of nonstopping to range from about 28 percent to 65 percent, and the finding of 56.9 percent in this study falls within this range. It appears, however, that the percentage of nonstopping vehicles varies from city to city and may have changed in recent years.

It should also be mentioned that conflict data were not collected relative to stopping characteristics of RTOR vehicles. The conflicts resulting from RTOR are highly dependent on pedestrian volumes, RTOR volume, side-street volume, and numerous locational factors. Thus a direct comparison of conflicts is not appropriate between RTOR-allowed and RTOR-prohibited sites, because sites may differ greatly in terms of pedestrian volume, RTOR volume, and so forth. It is possible, however, that a conflict problem on an intersection approach may exist because of the failure of RTOR vehicles to make a full stop. The magnitude of this RTOR conflict problem can only be determined on the basis of stopping characteristics data and corresponding conflict data at a large number of sites with RTOR allowed (i.e., 100 or more) with a variety of site and volume conditions.

Locational Factors Related to RTOR Violations

The next phase of the study involved determining geometric, traffic control, and other locational characteristics that are associated with high RTOR violation rates. The basic analysis approach for determining such related factors involved a safety engineering study of individual sites. This first involved ranking approaches by violation rate and then identifying common locational factors associated with high and low violation sites. This ranking was generated first for the 110 sites with RTOR prohibition, and then a separate ranking was developed of the 29 RTOR-allowed sites. These two situations are discussed in the following sections.

Locational Factors for RTOR-Prohibited Sites

Violation rates (turning right on red) at RTOR-Prohibited sites ranged from 0 to 25.6 percent. A distribution of the violation rates of the 110 sites was as follows:

Violations (%)	No. of Sites
0-1	13
1-2	21
2-3	19
3-4	11
4-5	6
5-6	11
6-8	7
8-10	4
10-12	7
12-18	8
18-30	3

The top 29 sites (26.3 percent) were found to have a violation rate greater than 6.0 and were labeled as the high-violation group. A total of 34 sites (30.9 percent) had a violation rate of 2 percent or less and were labeled as the low-violation group.

For the locations in the high- and low-violation groups, factors were identified that were related to high and low violations based on field inspections, a review of site diagrams, and a review of computer summaries of traffic data, signal data, and other information at each site. Location factors were identified as related to high violations if they were routinely found in the high-violation group but not in the low-violation group.

Traffic and roadway factors found to be typically associated with high violation rates include the following variables (individually or in various combinations):

1. Confusing or inappropriate partial prohibition signs [i.e., NTOR-SCHOOL DAYS ONLY sign located near a university, because motorists are not sure whether classes are in session on Saturdays, during summer sessions, and so forth; another NTOR sign near an elementary school prohibited RTOR during times after children had already arrived at school (9:00 a.m. to 2:00 p.m.) and ended before children left for home in the afternoon];
2. NTOR signs that are located on the far side or are inconspicuous to the motorists, particularly when placed on the far side across wide streets;
3. Combinations of low cross-street volume and low pedestrian volumes;
4. Approaches with easy right-turn maneuvers or right turns less than 90 degrees such as at Y-intersections, particularly with low conflicting movements;
5. Long cycle lengths that result in excessive waiting time for right-turn motorists;
6. High-speed ramps that form a T-intersection with a low-volume cross street;
7. Wide one-way streets on the cross street with low volume in the curb lane;
8. Confusing, multileg intersection approaches or approaches with an offset cross street;
9. Approaches where RTOR prohibition does not appear to be justified for some or all periods of the day because of low traffic volumes and little or no pedestrian traffic; and
10. Low right-turn volume per hour. [However, this is somewhat misleading because the percentage of violations is the total RTOR vehicles divided by the right-turn vehicles (including RTOG). As right-turn volume increases, a higher percentage of right-turn vehicles are trapped second, third, or fourth in line and cannot physically make an RTOR.]

The intersection approaches with low RTOR violation rates were also studied to determine related factors. The factors typically found at low-violation sites included the following variables:

1. Double NTOR signs located on the near and far sides, or NTOR signs that were located overhead or in a conspicuous location for stopped motorists;
2. High pedestrian volumes in either the near or far crosswalk (reduced opportunity for an RTOR);
3. High cross-street volume (reduced number of gaps and lower opportunity for an RTOR);
4. Crosswalk set back from the intersection farther than normal, combined with high pedestrian volumes;
5. Short signal cycle length;
6. A sharp right-turn maneuver (greater than 90 degrees) combined with poor sight distance;
7. High right turns per hour (however, this is misleading, as discussed previously); and
8. A cross street with on-street parking on the right, which forces an RTOR vehicle to make a wide turn beyond parked cars.

These results appear to indicate that motorist violations to NTOR signs are high when the signs are obscure or when it is not obvious to the driver why RTOR is prohibited (i.e., low pedestrian and cross-street volume and good sight distance). Drivers are particularly likely to run an NTOR sign at sites with long cycle lengths (when waiting time may be long). Some of the factors in the previous list were found to be useful for developing countermeasures.

Consideration was given to conducting more formal statistical analysis techniques to further support the factors that are associated with high and low violation rates. A branching analysis was conducted to identify roadway variables (independent variables) that account for the largest amount of explained variance in the violation rate (dependent variable). In addition to the branching analysis, preliminary Pearson correlation analysis and analysis of variance (ANOVA) tests were conducted. However, correlation coefficients were low (less than 0.3) for individual variables, and the ANOVA test required a larger data base of approaches to control for the interaction of traffic and roadway variables as they affect RTOR violation rates. It was evident that an engineering analysis of each approach was most useful in determining individual factors or combinations of factors that were related to high or low violation rates.

Locational Factors for RTOR-Permitted Sites

A detailed study was also made of traffic, geometric, and other factors at each of the 29 RTOR-permitted approaches to identify factors related to stopping violations (i.e., not making a full stop before turning right on red). At the 29 signalized approaches with RTOR permitted, no-stops ranged from zero to 45.2 percent, and total stopping violations (no stops plus rolling stops) ranged from 21.2 to 88.9 percent. One approach that had a sign posted RIGHT TURN ON RED ALLOWED AFTER STOP experienced 26.7 percent no-stops and 68.6 percent total stopping violations, compared with an overall average of the 29 sites of 14.8 percent no-stops and 56.9 percent total violations. It is possible that the sign had an effect of increasing stopping violations at the site, although insufficient data existed to verify this.

Locational factors found to be associated with a high rate of stopping violations included

1. Good sight distance with low pedestrian volume and low cross-street volume;
2. High right-turn volume;
3. Low pedestrian volume;
4. Low cross-street volume;
5. Unusual signal timing, such as split phasing, which minimized or eliminated conflicting traffic for part of the red interval;
6. Offset cross street (which lowered or delayed conflicting traffic and increased the opportunity for an RTOR rolling stop or no-stop); and
7. Nearby signalized intersection on the cross-street upstream, which created artificial gaps in cross-street traffic and provided greater opportunities for RTOR rolling stops or no-stops.

The factors found to be associated with low stopping violations at RTOR-allowed approaches included

1. High cross-street volume;
2. Poor sight distance (i.e., on-street parking on the cross street to the left of the approaching right-turn motorists);
3. High speed of cross street; and
4. High pedestrian volume.

These results indicate that drivers were more likely to comply with the stopping requirement when forced to do so (i.e., high pedestrian volume or cross-street traffic). Also, poor sight distance was a factor associated with high compliance, because drivers often made a full stop to look for cross-street traffic. During intervals of little or no pedestrian or conflicting traffic (such as with special signal phasing), motorists were less likely to make a full stop before turning right on red.

More formal statistical analysis techniques were not used for identifying related factors, because such analyses are not particularly appropriate for relatively small sample sizes of this type. The factors in the previous list were considered for development of possible countermeasures relative to RTOR stopping violations, as discussed in the next section.

SELECTION OF CANDIDATE COUNTERMEASURES

The factors related to high and low RTOR violations were studied and then grouped into corresponding high- and low-violation categories (Table 7). For example, one of the factors related to high violation of NTOR signs was long cycle length (excessive delay to right-turn motorist). A corresponding factor related to low violation rates was short cycle length. Thus, by grouping these factors, candidate countermeasures were developed, such as improving signal timing or installing traffic actuation devices.

As noted in Table 7, seven basic situations were found for which countermeasures could be proposed. Four of these situations related to violations of RTOR prohibitions and three involved the incidence of stopping violations (vehicles not making a full stop before an RTOR maneuver) where RTOR is allowed. For several of the violation causes, countermeasures were suggested that either may have an effect on the violation rates or may reduce the degree of hazard resulting from the violations. For example, for RTOR violations that involve not making full stops before turning right on red, countermeasures that may reduce the danger of such violations may include

1. Relocating the crosswalk farther from the intersection,
2. Warning pedestrians of possible right-turn danger through the use of WALK WITH CARE pedestrian

TABLE 7 Summary of Development of Candidate Countermeasures Based on Factors Related to RTOR Violations

High/Low Situation	Type of Violation Problem	Factors Related to High RTOR Violations	Factors Related to Low RTOR Violations	Candidate Countermeasures
1	RTOR where prohibited	NTOR signs located on far side or inconspicuous to the motorist	Double NTOR signs located on near and far side, or NTOR signs that are located overhead or in a conspicuous location for stopped motorists	1. Illuminate NTOR sign 2. Increase sign size to improve visibility 3. Relocate signs to near signal placement 4. Use double NTOR signs for redundancy 5. Use NTOR signs with red ball 6. Advanced warning of NTOR 7. Remove roadside clutter (to make NTOR sign more conspicuous) 8. Provide or improve intersection lighting
2	RTOR where prohibited	Confusing or inappropriate partial prohibition signing	Clear and visible NTOR signing	1. Prohibit RTOR only during the hours of heavy pedestrian travel 2. Use full RTOR prohibition on the approach 3. Use variable message NTOR signs 4. NTOR illuminated signal to be activated only during periods when RTOR is prohibited
3	RTOR where prohibited	Long cycle lengths (excess waiting time for right-turn motorists)	Short signal cycle lengths	1. Improve pedestrian signal display 2. Retime the traffic signal to provide better operations 3. Install presence detectors at traffic-actuated approaches to provide more efficient signal operation 4. Remove unwarranted traffic signals
4	RTOR where prohibited	Easy right-turn maneuver	Crosswalk set back from intersection farther than normal combined with high pedestrian volumes	1. Relocate crosswalk 2. Offset or angled stop bar 3. Special pavement marking in crosswalk
5	Stopping violations where RTOR allowed	Unusual signal timing	Lack of opportunity because of consistent traffic flow on cross street	1. Install flashing red right turning arrow to encourage full stop 2. Install NTOR sign if warranted 3. Retime traffic signal 4. Install part-time RTOR prohibition sign or variable message NTOR display 5. Install RIGHT TURN ON RED AFTER STOP sign to encourage full stops 6. Use special pedestrian signal display (i.e., WALK WITH CARE signal message during the WALK interval) 7. Install special pavement markings in crosswalk (i.e., LOOK FOR TURNING VEHICLES)
6	Stopping violations where RTOR allowed	Good sight distance	Poor sight distance	1. Install RIGHT TURN ON RED AFTER STOP sign to encourage full stops 2. Install YIELD TO PEDESTRIAN sign 3. Relocate crosswalk farther from intersection
7	Stopping violations where RTOR allowed	High right-turn volume, low pedestrian volume, or low cross-street volume	Low right-turn volume, high pedestrian volume, or high cross-street volume (or speed)	1. Install RIGHT TURN ON RED AFTER STOP sign to encourage full stops 2. Install NTOR sign if warranted 3. Install part-time RTOR-prohibition sign or variable-message NTOR display 4. Install YIELD TO PEDESTRIAN sign 5. Install PEDESTRIANS WATCH FOR TURNING VEHICLES sign 6. Use special pedestrian signal display (i.e., WALK WITH CARE signal message during the WALK interval) 7. Retime traffic signal 8. Remove unwarranted traffic signals 9. Relocate crosswalk further from intersection 10. Use special pavement marking in crosswalk (i.e., LOOK FOR TURNING VEHICLES) 11. Construct pedestrian overpass or underpass 12. Construct separate right-turn lane

Note: The countermeasures in this table were intended to correspond to traffic engineering treatments (i.e., improvement of traffic control devices or transportation facilities). It is recognized that provision of selective police enforcement and use of public education programs may also be of considerable benefit with respect to improving compliance and understanding or both of RTOR requirements and devices.

signals or LOOK FOR TURNING VEHICLES pavement markings, and

3. Constructing a pedestrian overpass or underpass to physically separate pedestrians and motorists.

Although RTOR motorists should yield to pedestrians, pedestrians should also be alert whenever crossing the street, because the pedestrian is usually the one who is injured in the event of a vehicle-pedestrian accident. Thus some of the countermeasures listed in Table 7 are intended to reduce violations related to RTOR, and other countermeasures are intended to reduce the potential hazard of RTOR maneuvers (either legal or illegal).

Based on all of the sources discussed previously, 30 potential RTOR-related accident countermeasures were devised (Table 8). These were categorized as they relate to signs, signals, pavement markings, design treatments, or other types of countermeasures.

For each countermeasure, a description is given along with comments and an indication regarding whether the countermeasure was selected for field testing. Many of these countermeasures may relate not only to RTOR and RTOR-pedestrian accidents, but to pedestrian accidents in general. A few of the countermeasures (i.e., eliminating unwarranted signals and retiming signals) may also affect other types of accidents (rear end, right angle, and so forth) and intersection operations (delay, congestion).

SUMMARY OF FINDINGS AND CONCLUSIONS

The purpose of this analysis was to conduct observational studies at signalized intersections to deter-

TABLE 8 Countermeasures Developed for RTOR

Category	Device	Description	Selected for Field Study	Comments
Signing	Full prohibition of RTOR	Install NTOR sign at locations with high traffic or pedestrian volumes, poor sight distances, at school crossings, or where other such factors influence the safe RTOR maneuver	No	There are some locations where RTOR maneuvers are unduly hazardous; although the MUTCD has guidelines on the application of NTOR signs, they are general and prone to a wide variety of interpretations; this leads to a nonuniform application of RTOR prohibitions; because conditions may change based on time of day, day of week, and season, a full-time prohibition may not always be warranted at a site
	Partial prohibition of RTOR for certain lanes or during specific times of the day	Install special signs that prohibit RTOR for certain times (7:00 a.m. to 7:00 p.m.), days (school days), conditions (when children are present), seasons (September to June), lanes (NTOR, except curb lane), or other factors	Yes	Because conditions may change at a site (by time of day or day of week), the prohibition should ideally only cover those times and conditions where warranted; however, some of the legends may require special knowledge by the motorists (school days), require motorists to drive "with one eye on the clock," or may be difficult to read
	YIELD TO PEDESTRIAN sign	Install a yield sign directed at turning motorists advising them to yield right-of-way to pedestrians	No	This device was tested in a previous FHWA study on pedestrian signalization alternatives and was found to be effective in reducing total right-turn conflicts with pedestrians
	Illuminate NTOR sign	Illuminate the NTOR sign for increased visibility; this could be accomplished by using an illuminated case sign (internal source) or external lighting	No	Designed for areas where there is a nighttime RTOR-related problem or where no intersection lighting exists or both
	Larger NTOR sign	Use an NTOR larger than the current MUTCD standard of 24 x 30 in. or 24 x 24 in.	Yes	NTOR sign should ideally be placed near the signal; it is applicable for near signal placement when the signal is located on the far side of a wide street or is otherwise difficult to read; it may be particularly helpful in cities or locations where overhead sign placement is not possible
	Near-signal placement of NTOR sign	Install NTOR sign on span arm, span wire, or signal pole near the signal head where motorist tends to look	No	MUTCD guidelines for NTOR sign placement state that signs should be located adjacent to the signal face to which they apply; many communities do not follow these guidelines and have the sign post mounted at the corner of the intersection
	Redundant NTOR signs	Install two or more NTOR signs on both posts (near or far side) and overhead to increase visibility of sign	No	Although this countermeasure is applicable for some locations with high violation rates, high conflict rates, or poor sign visibility, redundant sign placement should be minimized
	RIGHT TURN ON RED AFTER STOP sign	Install a sign that reminds motorist to come to a complete stop before turning on red	No	This device is intended to remind the driver to come to a full stop before making the RTOR maneuver, or to encourage more RTOR maneuvers where motorists are hesitant (and there are no conflicting pedestrian crossings or cross-street traffic)
	NTOR sign with red ball	Install a modified NTOR sign with a red ball in the center to draw attention to the sign	Yes	A sign with a red ball may catch the motorist's eye better; this device is currently used in some cities
	Advance warning of NTOR	Install a sign in advance of the intersection to warn motorists that there is an RTOR prohibition at the next intersection	No	This allows advance warning of conditions at the intersection and is consistent with positive guidance concepts; this sign may only add to the visual clutter of the roadside and may have minimal effect for those stopped at the signal

TABLE 8 Continued

Category	Device	Description	Selected for Field Study	Comments
Signing, continued	Electrical or mechanical variable message NTOR sign	Install signs that can display different messages for different signal intervals, times of day, or days of week	Yes	This device has two applications: (a) prohibit RTOR during portions of the day that have high pedestrian volumes or cross-street volumes, or (b) prohibit RTOR during portions of a cycle where a protected movement may conflict with the RTOR (such as an opposing protected left-turn maneuver); a blank-out display would avoid confusion when the message is not needed or other safety messages could be displayed; the cost for this device is expected to be high
	PEDESTRIANS WATCH FOR TURNING VEHICLES warning sign	Install a warning sign directed toward pedestrians to warn of turning vehicles; this device supplements pedestrian signals	No	This sign will not affect motorist behavior and is only applicable to pedestrians crossing the street; this may lead to additional visual clutter and is not effective for small children who cannot read; this device was tested in a previous FHWA study on pedestrian signalization alternatives and was found to be effective in reducing right-turn conflicts
Signals	Special pedestrian signal display (WALK WITH CARE)	Use a three-head signal that has a WITH CARE or other indication in yellow displayed during the WALK interval to warn of possible conflicts (i.e., WALK WITH CARE)	No	Special signal indications can be provided to remind the pedestrians to watch for turning vehicles; this type of device should only be used at locations where a known or potentially hazardous pedestrian problem exists, because overuse of such device could result in reduced effectiveness; this device was tested in a previous FHWA study on pedestrian signalization alternatives and was found to be effective in reducing right-turn pedestrian conflicts
	Retime traffic signal	Retime signal to reduce the conflicts and minimize delay; options include improved timing to accommodate flows, special pedestrian phasing, or use of multiphase operation	No	This is applicable to locations with high volumes of vehicle and pedestrian traffic, where turning movements are high, and where congestion is a problem; exclusive pedestrian crossing intervals, which have been noted to be related to lower pedestrain accidents, also increase delay and congestion to pedestrians and motorists
	Traffic-actuated signal	Use presence detectors to determine the right-turn demand and actuated signals to accommodate the demand and reduce the number of RTORs	No	May be applicable to some intersections with heavy right-turn demand
	Remove unwarranted traffic signals	Remove unwarranted signals and replace with other types of traffic control	No	Motorists lose respect for unwarranted signals, thereby increasing violations; many communities have begun programs to remove unwarranted signals where they no longer meet the warrants; although this may have the benefit of improving flow, reducing operating costs, and saving energy, pedestrians must cross the street without signal assistance
	Flashing red right-turn arrow	Install a flashing right-turn arrow to encourage motorists to come to a full stop before turning right on red	No	The flashing red arrow has been used in the past for right- and left-turn-on-red situations to stress the need for stopping before making an RTOR; this would require an extra signal lens; it may not convey a clear and simple meaning to all motorists and would require FHWA approval before use; it is currently not in the MUTCD
	NTOR signal installed in pedestrian signal hardware	Install an illuminated signal directed at motorists in pedestrian signal hardware to prohibit RTOR	No	This device uses existing pedestrian signal hardware (with a different lens) to display a blank-out or an NTOR indication to motorists; applicable for partial RTOR prohibitions; blank-out device minimizes confusion during RTOR-allowed periods
Pavement markings	Relocate crosswalk farther from intersection	Move the crosswalk farther from the intersection to increase visibility of pedestrians	No	Moving the stop bar and crosswalk farther from the intersection may discourage RTOR and increase the visibility of pedestrians; however, motorists failing to stop at the stop bar will block the crosswalk; this device may result in less sight distance of cross-street traffic and may encourage jaywalking
	Offset or angled stop bars	Angle or offset the stop bar so that drivers in the middle lanes are stopped farther back from the intersection than right-turn vehicles in the curb lane	Yes	For sites where RTOR is allowed; applicable to multilane approaches where there is a high incidence of truck and bus traffic that obstructs the drivers' view; allows the RTOR vehicle to see cross-street traffic and pedestrians for a safer turn; the effectiveness may be reduced if vehicles in the middle lanes do not observe the offset stop bar

TABLE 8 Continued

Category	Device	Description	Selected for Field Study	Comments
Pavement markings continued	Pavement marking	Pavement marking message in crosswalk to remind pedestrians to watch for RTOR vehicles (i.e., LOOK FOR TURNING VEHICLES)	Yes	The message is not visible to the motorist and will have no effect on driver reactions; installing pavement markings could create a slick surface for pedestrians unless a textured surface is used
Design	Pedestrian barriers	Install barriers to channelize pedestrians to the crosswalk, thereby minimizing the conflict area	No	The pedestrian barrier is also expected to reduce other types of pedestrian accidents, particularly dart-out and jaywalking-related accidents; however, barriers may cause difficulty in accessing parked vehicles along the curb, may be unsightly, and may create another roadside obstacle
	Pedestrian overpass or underpass	Grade separation of pedestrians and motorists to eliminate conflicts	No	Applicable to wide, high-speed intersections with safety problems; very expensive countermeasure, and the cost cannot be justified based on RTOR accidents alone; there may also be difficulties in accommodating elderly and handicapped pedestrians and bicyclists
	Far side bus stops	Allow buses to stop to drop-off and pick-up passengers only after crossing the intersection	No	Applicable where RTOR is allowed; eliminates congestion at the approach but may create a sight obstruction; far side bus stops are being used by many transit agencies to reduce intersection delays
	Eliminate parking near the intersection	Remove on-street parking near the intersection on either side or both sides of the street	No	On-street parking poses a site obstruction when near the crosswalk; this countermeasure may reduce other types of accidents at the intersection and may also increase capacity; however, it reduces parking availability; parking restrictions must be enforced to be effective
	Separate right-turn lane	Provide a separate lane for right turns and thus increase the opportunities for vehicle to make an RTOR	No	Applicable to sites with high volumes of right-turn traffic; increases the use of RTOR where RTOR is allowed; reduces intersection delay and increases capacity
Other	Intersection lighting	Illuminate the intersection to provide better visibility of pedestrians at night	No	Applicable to locations with high nighttime pedestrian volumes and where nighttime safety problems exist; may reduce other types of nighttime accidents at the intersection and may be useful in reducing crime at night
	Education campaign	Educate the public by using various forms of media to increase awareness and to teach proper understanding of RTOR	No	Educational campaigns can be directed at both the motorists and pedestrians related to RTOR safety and other safety issues; educational programs may not reach all individuals and may not have lasting impact; difficult to evaluate, especially relative to RTOR
	Clear roadside clutter	Remove roadside items to increase motorist visibility of pedestrians and traffic control devices	No	Removing all but essential roadside items should improve the motorist's ability to perceive pedestrians and traffic control devices and reduce distractions; may reduce other types of intersection accidents and improve aesthetics
	Selective traffic enforcement	Enforce violations of the NTOR sign and the requirement to complete a full stop before turning right on red where permitted; other pedestrian and motorist laws can also be enforced simultaneously	No	Enforcement or police presence near the intersection may reduce other violations; effectiveness may diminish once the police leave, because manpower is limited in most agencies; police time may be better spent in other areas of traffic enforcement or crime protection

mine current motorist compliance to RTOR prohibition and the requirement to make a full stop before turning right on red (where permitted). Traffic, geometric, and other physical site characteristics were collected in Detroit, Washington, D.C., and the Dallas/Austin area, and an in-depth engineering study was conducted at each of 110 intersection approaches where RTOR is prohibited. Data were also collected at 29 RTOR-allowed intersection approaches and 28 stop sign approaches relative to stopping characteristics (i.e., percentage of full stops, rolling stops, and no-stops of RTOR vehicles). Then locational factors were identified relative to high and low violation rates. The following is a summary of key findings and conclusions:

1. Overall, only 3.7 percent of all right-turning drivers violate the RTOR prohibition signs, based on a sample of more than 67,000 drivers. However, of those motorists given an opportunity to commit an RTOR violation, about 20 percent of them violate the NTOR sign.

2. Of the drivers who commit an RTOR violation, about 23.4 percent of them result in conflicts with pedestrians or cross-street traffic. However, less than 1 in 100 of the total right-turn vehicles is involved in an RTOR-related conflict.

3. At a sample of RTOR-prohibited sites, 22.7 percent of the illegal RTOR maneuvers resulted in a conflict with cross traffic or pedestrians. However, only 14.2 percent of RTOR maneuvers resulted in a

conflict to pedestrians, compared with 19.5 percent RTOG maneuvers that involve a pedestrian conflict.

4. Of the 29 intersection approaches with RTOR allowed, 26.2 percent of right-turn vehicles turned right on red. Of the vehicles turning right on red, the violation rate (not making a full stop) was 56.9 percent. This rate was higher for Washington, D.C. (61.4 percent of vehicles not fully stopping) and Detroit (59.1 percent), compared with Dallas/Austin (50.3 percent).

5. The overall violation rate (percent not fully stopping) at the 28 stop sign approaches was 68.2 percent compared with 56.9 percent for signalized approaches with RTOR allowed, a difference of 11.3 percent. However, 36 percent of vehicles were found to stop at RTOR-allowed approaches compared with 24.7 percent at stop sign locations. Thus the 11 percent higher violation rate at stop sign locations may be at least partly explained by the greater percentage of opportunities for a rolling stop or no-stop.

6. Examples of physical site factors found from in-depth site studies to be related to high RTOR violation rates include confusing or inappropriate partial prohibition signs; far side or inconspicuous NTOR signs; long cycle lengths; confusing multileg intersection approaches; unjustified RTOR prohibition; split-phasing of the signal, which creates low opposing traffic for RTOR maneuvers; and combinations of a low volume or high speed of cross-street traffic and low pedestrian volumes.

REFERENCES

1. Manual on Uniform Traffic Control Devices for Streets and Highways. FHWA, U.S. Department of Transportation, 1978.
2. M.R. Parker, Jr. The Impact of General Permissive Right- and Left-Turn-on-Red Legislation in Virginia. Report VHTRC 74-R7. Virginia Highway and Transportation Research Council, Charlottesville, Sept. 1978.
3. W.E. Baumgaertner. After Stop Compliance with Right-Turn-on-Red After Stop. ITE Journal, Vol. 51, No. 1, Jan. 1981.
4. W.A. McGee, W.A. Simpson, J. Cohen, G.F. King, and R.F. Morris. Right-Turn-on-Red. Report FHWA-RD-76-89. FHWA, U.S. Department of Transportation, May 1976.
5. D. Galin. Re-Evaluation of Accident Experience with Right-Turn-on-Red. ITE Journal, Vol. 51, No. 1, Jan. 1981.
6. T.D. Davis and W.L. Mullowney. Comparison of Right-Turn-on-Red and No-Turn-on-Red Traffic Performance. Presented at 63rd TRB Annual Meeting, Washington, D.C., Jan. 16-20, 1984.
7. M.R. Parker, Jr., R.F. Jordon, Jr., J.A. Spencer, M.D. Beale, and L.M. Goodall. Right-Turn-on-Red: A Report to the Governor and General Assembly of Virginia. Report VHTRC 76-R9. Virginia Highway and Transportation Research Council, Charlottesville, Sept. 1975.
8. Massachusetts Executive Office of Transportation and Construction. Right-Turn-on-Red Safety Study for Massachusetts. Governor's Energy Policy Office, Boston, April 28, 1978.
9. De Leuw, Cather and Company. Effect of Control Devices on Traffic Operations--Interim Report. NCHRP Report 11. HRB, National Research Council, Washington, D.C., 1964, 107 pp.
10. R.F. Baubien. Stop Signs for Speed Control? Traffic Engineering Magazine, Nov. 1976.
11. R.J. Benke and G.L. Ries. Right-Turn-on-Red: Permissive Signing Versus Basic Law. Minnesota Department of Transportation, St. Paul, Aug. 1973.
12. M.S. Mamlouk. Right-Turn-on-Red: Utilization and Impact. Report JHRP 76-17. Joint Highway Research Program, Purdue University, West Lafayette, Ind., June 1976.

Publication of this paper sponsored by Committee on Pedestrians.

Volume Guidelines for Signalization of Diamond Interchanges

MYUNG-SOON CHANG and CARROLL J. MESSER

ABSTRACT

The objective of the work described in this paper is to establish volume guidelines for the installation of traffic signal control at diamond interchanges where the base condition is all-way stop sign control. The guidelines are based on operational threshold values of traffic flow, above which signalization is expected to produce superior performance. Four diamond interchanges were studied with both types of control, from which the study results were based. The data-collection methods and procedures employed in the study to evaluate the operational effects of stop sign and signal control at diamond interchanges are discussed. An assessment of traffic control alternatives is described in terms of operational effects of queues and travel speed. Guidelines for all-way stop signs or signal control at diamond interchanges are provided in terms of internal volume, left-turn proportion within internal volume, and the sum of internal and external volume. The specific traffic volume guidelines were developed based on a combination of these variables, which affect operational performance.

Diamond interchanges are widely used in urban areas as a means to transfer freeway traffic to and from the surface street system. The selection of the proper traffic control system for each diamond interchange is a challenging task. When and where to use stop signs or signals for traffic control at a significant number of diamond interchanges is a principal concern. This complex subject is discussed in this paper and useful information is provided for guiding future engineering decisions in the selection of the appropriate diamond interchange control.

Signalization of a diamond interchange is often resorted to after public pressure is applied and one or both sides of the interchange are warranted by Manual on Uniform Traffic Control Device (MUTCD) (1) standards for a single intersection. However, MUTCD warrants for signalization neither explicitly reflect the operational characteristics of diamond interchanges nor are they sensitive to the traffic patterns associated with the two intersections at a diamond interchange.

Research conducted by the Texas Transportation Institute (TTI) regarding the operational characteristics of diamond interchange controllers led to a better understanding of different phasing patterns, and the development of frontage road progression strategies and a diamond interchange signal optimization and analysis program for timing pretimed diamond interchanges (2-4). FHWA also sponsored a series of research studies on signalized diamond interchanges, with particular emphasis on signal phasing (5-7).

The MUTCD provides national standards for determining when a signal is warranted at an intersection. The Texas manual (8) includes all eight MUTCD warrants plus an actuated control warrant. However, neither manual specifically considers diamond interchanges and their special requirements. One case study of a diamond interchange in Texas (9) illustrated a signal warranting situation where one side of an interchange was warranted and the other fell short. It was noted in this study that current signal warrant conditions do not appear to adequately address the different traffic movement patterns associated with two intersections at a diamond interchange.

The development of clear and effective guidelines for installing all-way stop signs or signals for traffic control at a significant number of diamond interchanges, whose traffic patterns and geometric physical characteristics vary quite widely between interchanges, would be a significant contribution to the traffic engineering technology.

The objectives of this study were as follows: (a) conduct an operational evaluation of the two types of traffic control (i.e., all-way stop and traffic signals) to include comparisons of vehicular delay and stops at diamond interchanges under various types of geometric and traffic patterns, (b) analyze operational results to determine the relative efficiency of each type of control, and (c) develop guidelines to aid in the selection of the appropriate control method for isolated interchanges.

EXPERIMENTAL PLAN AND ANALYSIS APPROACH

Type of Control

An experimental plan was developed to field evaluate the operational performance of two types of diamond interchange control strategies: all-way stop sign control, and traffic signal control. To provide a general guideline for signal control, signal operations were confined neither to a single controller type nor to a single phase pattern. Signal control in this study encompassed pretimed control, actuated control, three-phase operation, and four-phase overlap operation.

Study Sites

Field studies were conducted to evaluate the operational performance of stop sign and signal control. Four sites were selected for this study. The sites

TABLE 1 Operational Performance Data Collected at Study Sites

Interchange Location	Traffic Control Studied	Queue Counts: Stop Sign	Queue Counts: Signal	Travel Time: Stop Sign	Travel Time: Signal
US-83 at South 7th in Abilene	Stop sign, three-phase operation, four-phase overlap	X	X	X	X
US-59 at Jetero Boulevard in Houston	Stop sign, four-phase overlap	X	X	NA	NA
I-10 at T. C. Jester in Houston	Stop sign, four-phase overlap	NA	NA	X	X
I-20 at Trail Lake in Fort Worth	Stop sign, four-phase overlap	NA	NA	X	X

Note: NA = not available.

were selected to provide a variety of geometric and traffic conditions.

Data on the locations of the four sites and the overall field data-collection effort, as conducted, are given in Table 1. A wide variety of geometrics, traffic volumes, and traffic patterns was provided by the four sites. Two interchanges were underpasses and the other two interchanges were overpasses. Separation between intersections ranged from 260 to 480 ft. The number of lanes for each approach at the four interchanges ranged from one to three.

Besides all being located in major Texas cities, there were some other similarities in the four sites. All frontage roads were continuous through the interchanges without any U-turn lanes. All interchanges studied, except I-20 at Trail Lake, had left-turn bays between the two intersections.

Traffic control was varied among the interchanges. Some interchanges had a protective left-turn-only phase, whereas others had protective and permissive left-turn phases. Except at Abilene, stop sign performance was observed before signal installation. For Abilene, signal control was converted to stop sign control for a day, and the performance was observed the next day. All pretimed signals were operated at a 60-sec cycle length. The signal at I-20 at Trail Lake was the only actuated signal observed. Neither interchange design features nor signal control promoted highly efficient signal operations.

The study plan called for data to be collected for 4 hr per day from 7:00 to 8:00 a.m., 10:00 to 11:00 a.m., 12:00 to 1:00 p.m., and 5:00 to 6:00 p.m., or some reasonable on-site modification if deemed appropriate.

Several types of performance data were to be collected. The initial plan called for tracing vehicles through the interchange to obtain their travel time or travel speed along with their stopped delay. This was performed by recording an arrival time to the interchange influence zone, stopping times at Intersections 1 and 2, and departure times at Intersections 1 and 2. The count of the number of stopped vehicles on each approach was added later. The data in Table 1 give the performance data collected for alternative traffic controls at each interchange.

Traffic volumes were collected manually or by using automatic counters on all four inbound approaches to the interchange and on both interior intersection approaches. Two people, one for each intersection, were used to manually count traffic volume. Each approach flow was obtained for 15-min time periods and expanded to an equivalent hourly volume.

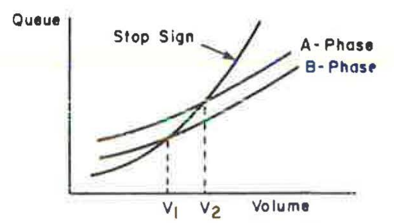
Additional manual observations were made every 15 sec during the study by six persons to determine the number of vehicles stopped on each of the six intersection approaches. Stopped vehicle data were recorded on scribble pads and then later reduced in the office. A 15-min time interval was used as the time base for data analysis.

The study supervisor observed general characteristics of traffic flow on the cross street and ramp traffic. Particular attention was paid to the effect of internal volume and its left-turn volume on traffic flow at an interchange.

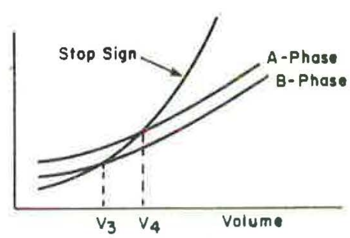
Analysis Approach

To provide guidelines for traffic control alternatives at diamond interchanges, the following three methods appear to be relevant:

1. Provide guidelines by separate signal control methods:

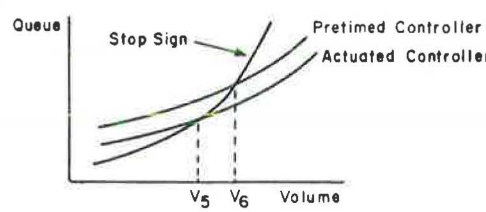


PRETIMED CONTROLLER

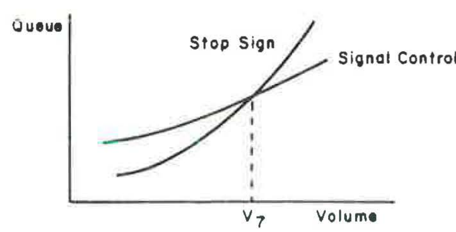


ACTUATED CONTROLLER

2. Provide guidelines by controller types:



3. Provide guidelines by general control alternatives:

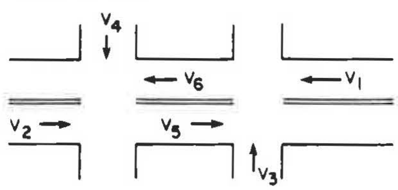


Because the objective of the study was to provide general guidelines for stop sign versus signal control, the third method was used throughout the study. However, every effort was made to distinguish performance differences between stop sign and signal control because of different interchange geometric and traffic characteristics.

Approach Used to Develop Guidelines

It is emphasized that guidelines should distinguish different geometric and traffic characteristics between different interchanges. The traffic volume on each approach was normalized with respect to approach lanes (i.e., the traffic volume on each approach was divided by its number of lanes) to distinguish geometric differences in the number of lanes on each approach among different interchanges. Thus the total interchange hourly volume per lane, which is the basic interchange volume used throughout this paper, was defined as the sum of the six intersection approach volumes per lane. Further, to distinguish different traffic patterns among different interchanges, two variables that characterize diamond interchange traffic movement were introduced:

1. Ratio of internal volume per lane to external volume per lane (RIE):



RIE = (Internal volume per lane)/(External volume per lane) = $(V_5 + V_6)/(V_1 + V_2 + V_3 + V_4)$.

The RIE variable reflects observations that stop sign control causes more delay to internal traffic and, subsequently, to overall interchange traffic than does signal control. Stop sign control requires double stops for all external volumes that use both intersections, whereas signal control usually provides progression through the interchange.

2. Composition of left-turn and through volume within internal volume: The reason for distinguishing left-turn from through volume within the internal traffic is that as more traffic turns left within the internal stations, overall interchange operation appears to be affected. Another reason for this distinction is to reflect the advantages and disadvantages of U-turn lanes to accommodate double left-turning traffic coming from frontage roads.

STUDY RESULTS

A presentation of the results of the field studies follows. A general description of the traffic volumes, travel speeds, and queue characteristics observed at each diamond interchange will introduce the findings. Detailed statistical analyses to assess stop sign and signal control and their results by type of traffic control conclude this section.

Traffic Volumes

The data in Table 2 present the range of interchange traffic volumes observed at the four interchanges. The four interchanges are sequenced according to the rank of highest volume levels. Observed total interchange hourly volume per lane at the four interchanges ranged between 600 and 2,000 vehicles.

TABLE 2 Ranking of Four Interchanges by Observed Total Interchange Hourly Volume per Lane

		Volume	
Rank	Interchange Location	Highest	Lowest
1	US-59 at Jetero Boulevard in Houston	1,999	692
2	I-20 at Trail Lake in Fort Worth	1,773	889
3	US-83 at South 7th in Abilene	1,658	886
4	I-10 at T. C. Jester in Houston	855	607

Travel Speeds

Travel times were traced at each of the four external stations at each interchange. The reference point from which traffic is assumed to be influenced by traffic control (stop sign or signal) was established as a utility pole or sign pole located approximately 300 to 500 ft away from the stopline on each approach. When a vehicle passed the reference point, its time was recorded. The vehicle was traced with regard to its travel time and direction of movement until it was completely out of the interchange. The stop delay is the sum of the differences between the departure time and stop time at an intersection within the interchange. Travel time is the difference in time between arrival time to the outer reference point and the departure time from the last intersection.

To normalize the differences in distances traveled by a vehicle at each interchange, all travel times were converted to travel speeds. Further, those directional movements passing through two intersections were distinguished to reflect the diamond interchange characteristics. In addition, through and left-turn movements were separated because their speeds appeared to be affected differently by the traffic control alternatives.

Travel speeds involving left-turning vehicles, observed at the four interchanges, ranged from 26.9 to 4.4 ft/sec for stop sign control, and from 29.1 to 5.0 ft/sec for signal control. For cross-street through traffic, travel speeds observed ranged from 23.1 to 5.6 ft/sec for stop sign control, and from 29.4 to 6.2 ft/sec for signal control. Generally, travel speeds were observed to decrease as total interchange traffic volume increased.

Queue Characteristics

It was noted in the previous discussion that the number of stopped vehicles was observed at six interchange stations (or approaches). Two stations (Stations 1 and 2) were on the arterial cross street and another two stations (Stations 3 and 4) were located on the frontage roads. The remaining two stations (Stations 5 and 6) were located between the traffic signals. To account for the different number of traffic lanes on each approach, the number of stopped vehicles was divided by the number of lanes on each approach.

Therefore the total interchange queue is defined as the sum of the average number of vehicles observed to be stopped per lane at the six stations of the interchange. The traffic queue on an approach (station) is an average value across all lanes and is not a critical lane value. Queue counts were taken every 15 sec and averaged over 15-min intervals.

Overall, less queue was observed for stop sign control than signal control when interchange traffic volume was low. As interchange traffic increased, such as during peak hours, more queue was observed for stop sign control than for signal control. These general trends were observed for all interchanges studied.

Figure 1 shows the queue characteristics observed at the interchange in Abilene, Texas. It revealed the following characteristics:

1. As traffic volume increased, signal control was a more effective alternative in reducing queue than stop sign control, and

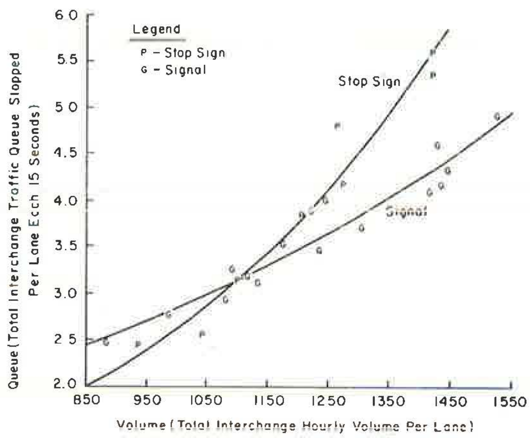


FIGURE 1 Queue versus volume by stop sign and signal control in Abilene.

2. As traffic volume increased to more than 1,100 per hour per lane, signal control was more effective than stop sign control.

Figure 2 shows the queue characteristics observed at the interchange in Houston, Texas. It revealed the following characteristics:

1. It confirmed the general expectations that as traffic volume increased, traffic signal control was more effective in reducing queue than stop sign control, and
2. As traffic increased beyond 600 vehicles per hour per lane, traffic signals were more effective than stop signs.

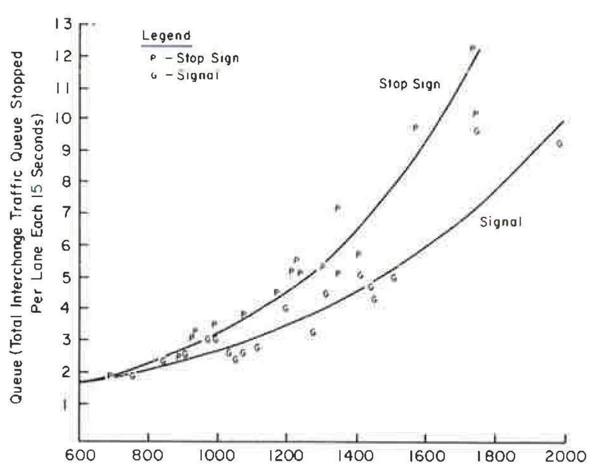


FIGURE 2 Queue versus volume by stop sign and signal control in Houston.

Comparing Figure 1 for Abilene with Figure 2 for Houston, it is noted that the intersecting point, which has approximately equal queue generation for both stop signs and signal controls, is different between the interchanges. These differences are caused in part by different interchange traffic patterns. This consequence is reflected in the development of guidelines on when and where a stop sign or traffic signal is preferred.

Assessment of Traffic Control Alternatives

The assessment of traffic control alternatives involves two areas. The first examines performance differences between stop sign and signals for their effects on queue. The second evaluates differences between stop sign and signals for their effects on travel speed and travel time. These two areas of interest initially will be analyzed separately. Later, the queue and travel speed information will be combined to suggest volume guidelines for signal control.

Relationship Between Queue and Volume by Traffic Control

The initial data analysis from Abilene and Houston revealed that when more traffic flows between the two intersections (such as left turns from the ramp and through traffic on the arterial), traffic signals are more effective at lower interchange volumes than in the case of traffic using only a single intersection (such as through traffic from ramps and right-turn traffic from arterials).

The queues observed from Abilene and Houston were pooled together. Two-dimensional plots of queue versus total interchange traffic volume per hour per lane indicated that an exponential function would fit the observed data well. Another variable that characterizes traffic movements that encompass two intersections between signals--the ratio of internal volume to external volume--was added. The exponential form used is as follows:

$$Q = \mathrm{Exp}(a + bV + c\mathrm{RIE})$$
$$Q = A\ \mathrm{Exp}(bV + c\mathrm{RIE}) \quad (1)$$

where

Q = total interchange traffic queue stopped per lane as observed each 15 sec,
V = total interchange traffic volume per hour per lane,
RIE = ratio of internal volume to external volume, and
A,a,b,c = derived coefficients.

The logarithm transformation of Equation 1 can be linearized as log Q = a + bV + cRIE. By using the Statistical Analysis System (SAS) (10), models for stop sign and signal control were derived. Models that describe the total number of stopped vehicles at interchange per lane were developed as follows:

Stop sign control: $Q_p = 0.26\ \mathrm{Exp}(1.89\ v/1000 + 0.94\ \mathrm{RIE}) \quad (2)$

Signal control: $Q_s = 0.29\ \mathrm{Exp}(1.25\ V/1000) \quad (3)$

The coefficients of determination (R^2) for stop sign and signal control were 0.95 and 0.93, respectively. All variables are significant at the α = 0.01 level. The RIE variable for signal control was not statistically significant (α = 0.25). Signal progression apparently handles substantial internal traffic more efficiently than stop sign control.

Plots of queue versus volume for stop sign and signal control are shown in Figure 3. The plot of stop sign control is represented by the typical ratio of internal volume to external volume observed in the field (i.e., four cases of RIE = 0.4, 0.5, 0.6, and 0.7). Note in Figure 3 that the faster more internal traffic occurs at an interchange (i.e., larger RIE), the sooner signal installation is needed.

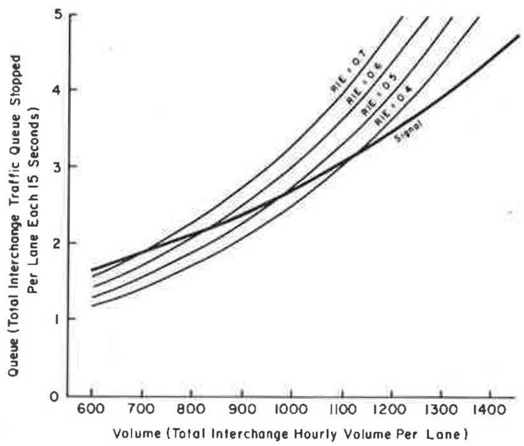


FIGURE 3 Queue versus volume by stop sign and signal control.

Specifically, the models and plots of queue performance revealed the preferences to the type of traffic control given in Table 3. Note in Table 3 that the diamond interchange should be considered as a special category different from intersections in which interchange operation is sensitive to the degree of internal traffic movements between the two signals.

TABLE 3 Traffic Control Alternative Performance Based on Queue Only as Related to Total Interchange Volume

	Volume	
RIE	Shorter Queue During Stop Sign Control	Shorter Queue During Traffic Signal Control
0.4	< 1,140	> 1,140
0.5	< 990	> 990
0.6	< 840	> 840
0.7	< 690	> 690

Note: Total interchange volume is the sum of internal and external traffic volume per hour per lane at an interchange.

Relationship Between Travel Speed and Volume by Traffic Control

Travel speed is analyzed by traffic movements because the travel speed for through movements on the cross street is different from traffic movements that involve left turns from cross streets and ramps. Further, it is hypothesized that travel time is affected by the degree of internal traffic at an interchange. The model used to evaluate the travel speed at an interchange was developed as follows.

For arterial through traffic movements:

Stop sign control: $U_p = 26.61 - 9.07\ V/1000$ (4)

Signal control: $U_s = 81.93\ \mathrm{Exp}(-0.53\ V/1000 - 1.62\ \mathrm{RIE})$ (5)

For left-turn traffic movements:

Stop sign control: $U_p = 28.93 - 10.17\ V/1000$ (6)

Signal control: $U_s = 39.66\ \mathrm{Exp}(-0.35\ V/1000 - 0.88\ \mathrm{RIE})$ (7)

where

U_p = travel speed for stop sign control (ft/sec),
U_s = travel speed for signal control (ft/sec),
V = total interchange traffic volume per hour per lane, and
RIE = ratio of internal traffic volume to external traffic volume.

Travel speed for stop sign control did not statistically depend on the degree of internal traffic movements. The reason appears to be that the relative stop delay for stop sign control is not sensitive enough because of its regularity by all approach traffic. However, travel speed for signal control is sensitive to internal traffic movements because they influence progression speed from the cross street and ramps.

Plots of travel speed versus volume for left-turn and arterial through traffic are shown in Figures 4 and 5, respectively. The model and plot of travel speed performance revealed the following:

1. For arterial through traffic, signalization appears to perform better than stop sign control

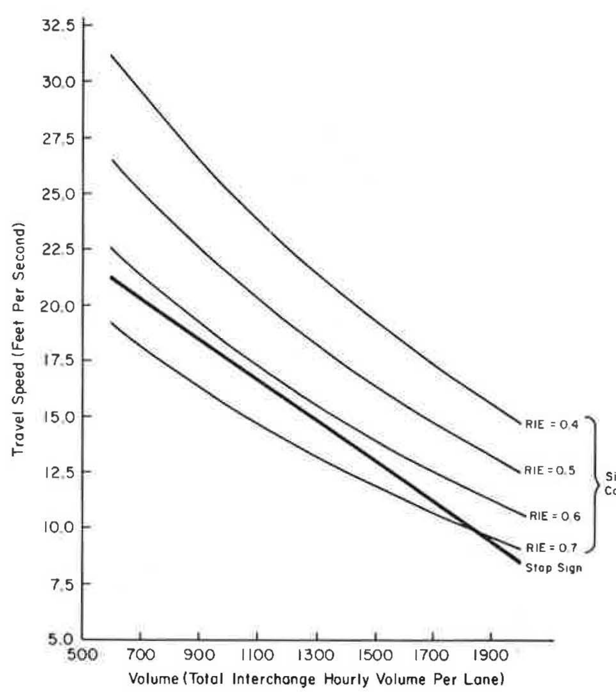


FIGURE 4 Travel speed versus volume for arterial through traffic by stop sign and signal control.

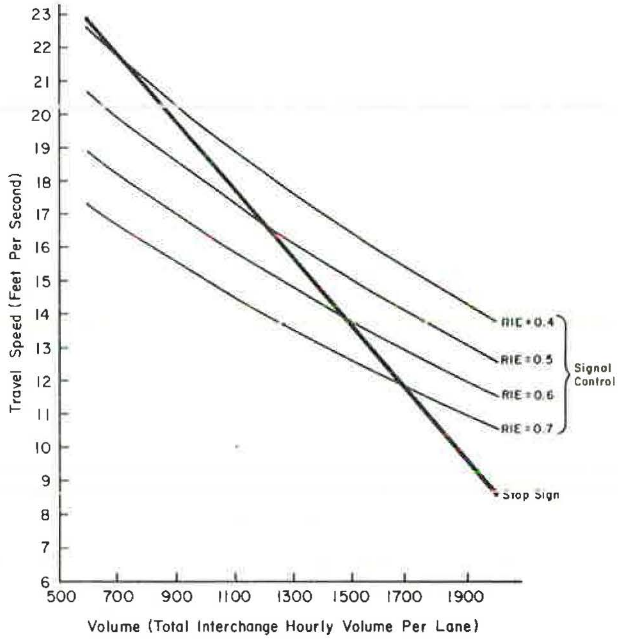


FIGURE 5 Travel speed versus volume for left-turn traffic by stop sign and signal control.

unless internal volume reaches 70 percent of external traffic. The reason appears to be that signal control can maintain relatively good progression until internal volume becomes substantial enough to affect external approach traffic.

2. For left-turning traffic, stop signs appear to perform better than signal control unless interchange traffic and internal traffic reach critical volume levels. The reason appears to be that left-turning traffic often has to wait a cycle with signal control, whereas stop sign control does not require this traffic to wait a cycle.

Development of Guidelines Combining Queue and Travel Speed Results

A sample problem is introduced to illustrate the procedure employed to develop volume guidelines of signal control considering the queue and travel speed findings. A complete set of guideline volumes will be presented after the sample problem illustration.

Assume an interchange has an RIE (i.e., the ratio of internal volume over external volume) equal to 0.50. The volume guideline for signalization at this interchange would be 990 vehicles per hour per lane if queue were the only measure of effectiveness considered (see Figure 3 and Table 3).

Considering travel speed or travel time, signals are more efficient for arterial through traffic, but stop signs are more efficient for left-turning traffic at this volume level (see Figures 4 and 5).

The adjustment procedure for travel speed is as follows. Assume that 40 percent of internal traffic turns left and the other 60 percent goes through. The speed ratios observed between stop sign and signal control for left-turn and arterial through traffic are as follows:

For the left-turn speed ratio:

$$\text{Stop/Signal} = (28.93 - 10.17 \times \text{Volume}/1000) \div [39.66 \text{ Exp}(-0.35\ V/1000 - 0.88\ \text{RIE})]$$

$$= (28.93 - 10.17 \times 0.99)/[39.66 \text{ Exp} \times (-0.35 \times 0.99 - 0.88 \times 0.5)]$$

$$= 18.9/18.1 = 1.04 \quad (8)$$

For the arterial through traffic speed ratio:

$$\text{Stop/Signal} = (26.61 - 9.07\ V/1000)/[81.93 \text{ Exp} \times (-0.53\ V/1000 - 1.62\ \text{RIE})]$$

$$= (26.61 - 9.07 \times 0.99)/[81.93 \text{ Exp} \times (-0.53 \times 0.99 - 1.62 \times 0.5)]$$

$$= 17.6/21.6 = 0.81 \quad (9)$$

Because there is 40 percent left-turn traffic and 60 percent through traffic at this interchange, the adjustment ratio is

$$\text{Stop/Signal} = 40 \text{ percent} \times \text{left-turn ratio} + 60 \text{ percent} \times \text{through ratio}$$

$$= 0.4 \times 1.04 + 0.6 \times 0.81$$

$$= 0.90 \quad (10)$$

This means that a signal is more efficient than stop signs in travel speed for this traffic pattern. Specifically, signal control is 11 percent faster (i.e., 1/0.90 = 1.11) than stop sign control.

Considering this travel speed efficiency, traffic engineers would like to install a signal sooner than the 990 volume level. This means that an adjustment should be made to reflect travel speed efficiency in addition to queue considerations, as follows:

Guideline based on travel speed = 990 x 0.90 = 890 vehicles.

Figure 6 shows the adjustment effect based on travel speed. Assuming an equal weight between queue and travel speed performance, the guideline would be about 940 vehicles [i.e., (990 + 890)/2] in this example.

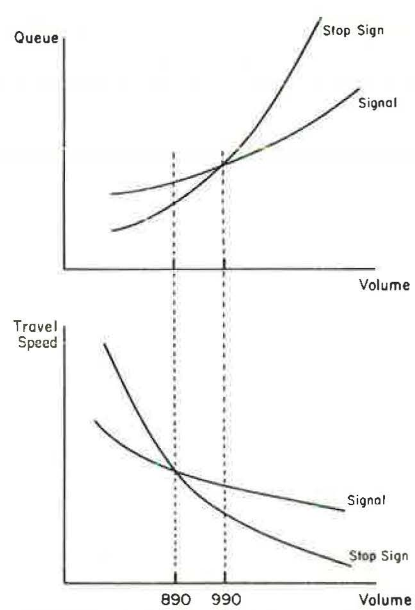


FIGURE 6 Adjustment effect of queue and travel speed.

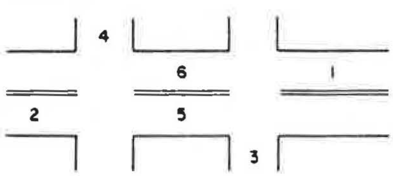
Signalization Guidelines

Following the procedure illustrated in the previous example, various combinations of internal traffic and left-turn traffic observed in the field were considered. RIEs from 0.4 to 0.7 were evaluated together with left-turn proportions from 30 to 70 percent. The results obtained are given in Table 4, which gives the recommended volume guidelines for installing signals at diamond interchanges.

TABLE 4 Guidelines for Installing Traffic Signals at Diamond Interchanges

RIE	Left Turn (%)	Minimum Interchange Volume for Signal Control
0.4	30	1,005
	50	1,035
	70	1,060
0.5	30	935
	50	955
	70	980
0.6	30	850
	50	865
	70	885
0.7	30	750
	50	760
	70	775

Note: RIE is the sum of internal traffic volume per hour per lane divided by the sum of external traffic volume per hour per lane; left turn (%) is the proportion of left-turn traffic within internal traffic; interchange volume for signal control is the sum of internal and external traffic per hour per lane at an interchange; internal traffic is traffic at Stations 5 and 6; and external traffic is traffic at Stations 1, 2, 3, and 4:



If the suggested guideline volumes presented in Table 4 are applied following MUTCD practice, then these volume levels must be exceeded for each of any 8 hr of an average day. However, the exact number of hours required to meet the guideline volume levels for implementation should be determined from further study and testing in practice.

Simplified Guidelines

It is noted in Table 4 that the interchange volume guidelines for signal control are practically insensitive to left-turn proportion within internal volume. Considering the effort required to collect the data, the left-turn proportion could be practically negligible for implementation. From these considerations, the simplified guidelines given in Table 5 are also provided for this practical reason.

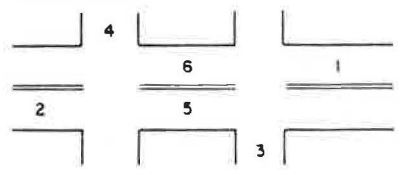
Comparison with MUTCD Warrants

The MUTCD states that traffic control signals should not be installed unless one of the signal warrants in the manual is met. Two of the warrants in the manual are related to traffic volume.

TABLE 5 Simplified Guidelines for Installing Traffic Signals at Diamond Interchanges

RIE	Minimum Interchange Volume for Signal Control
0.4	1,050
0.5	950
0.6	850
0.7	750

Note: RIE is the sum of internal traffic volume per hour per lane divided by the sum of external traffic volume per hour per lane; interchange volume for signal control is the sum of internal and external traffic per hour per lane at an interchange; internal traffic is traffic at Stations 5 and 6; and external traffic is traffic at Stations 1, 2, 3, and 4:



The first warrant--Minimum Vehicular Volume--is intended for application where the volume of intersecting traffic is the principal reason for signal installation. The warrant is satisfied when, for each of any 8 hr of an average day, the traffic volumes given in Table 6 exist on the major street and on the higher-volume minor street approach to the intersection.

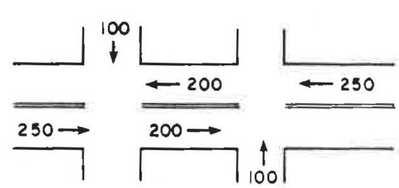
TABLE 6 MUTCD Minimum Vehicular Volumes for Warrant 1

No. of Lanes for Moving Traffic on Each Approach: Major Street	Minor Street	Vehicles per Hour on Major Street (total of both approaches)	Vehicles per Hour on Higher-Volume Minor Street Approaches (one direction only)
1	1	500	150
2+	1	600	150
2+	2+	600	200
1	2+	500	200

The second warrant--Interruption of Continuous Traffic--applies to operating conditions where the volume on the major street is so heavy that traffic on the minor intersecting street suffers excessive delay or hazard in entering or crossing the major street. Thus the second warrant is only applicable to two-way stop sign control. Therefore, the second warrant is not applicable to all-way stop sign control at diamond interchanges.

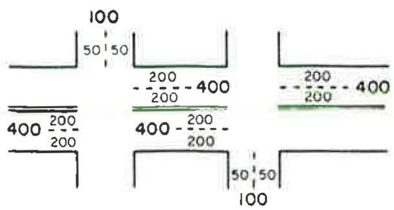
Examples are presented to compare the MUTCD warrant with the guidelines derived from this study (which are called diamond interchange guidelines).

1. Example 1: One lane for all approaches that have traffic volumes:



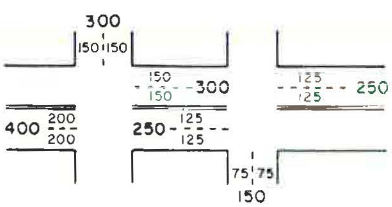
Because the major street carries 450 vehicles and the minor street carries 100 vehicles, neither intersection will satisfy MUTCD warrant 1. However, because the total interchange volume is 1,100 vehicles per lane at an internal ratio of 0.6, this sample interchange will satisfy the diamond interchange guidelines.

2. Example 2: Two lanes for all approaches that have traffic volumes:



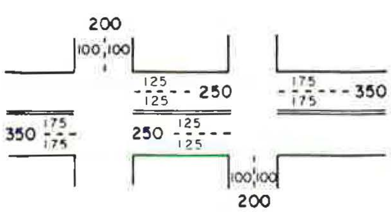
Because the major street carries 800 vehicles and the minor street carries 100 vehicles, neither intersection completely satisfies the warrant. However, because the total interchange volume is 900 vehicles per lane at an internal ratio of 0.8, this sample interchange will satisfy the diamond interchange guidelines.

3. Example 3: Unbalanced traffic flow:



Intersection 1 satisfies MUTCD warrant 1 but Intersection 2 does not. The option of installing two separate traffic controls (e.g., signals at Intersection 1 and stop signs at Intersection 2) at the interchange is too risky to use. Assume that signals are installed at this interchange because Intersection 1 warrants signalization. However, because the diamond interchange carries 825 vehicles per lane at an internal ratio of 0.5, this interchange will not satisfy the diamond interchange guidelines for signalization.

4. Example 4: MUTCD warrant is met but diamond interchange guidelines are not met:



Because the major street carries 600 vehicles and the minor street carries 200 vehicles, both intersections meet MUTCD warrant 1 for signalization. However, because the interchange carries 800 vehicles per lane at an internal ratio of 0.45, it does not meet the diamond interchange guidelines.

Numerous other examples can be illustrated in which the following four cases exist:

1. MUTCD warrant is met, but diamond interchange guidelines are not met;
2. MUTCD warrant is not met, but diamond interchange guidelines are met;
3. MUTCD warrant is met for one intersection and is not met for another intersection, but diamond interchange guidelines are met; and
4. MUTCD warrant is met for one intersection and is not met for another intersection, but diamond interchange guidelines are not met.

From these possible cases it is noted that the two intersections at a diamond interchange cannot be separated regarding their operational characteristics. The independent treatment of two intersections at a diamond interchange is improper. Thus diamond interchanges should be treated as a separate warrant category in the MUTCD. The interchange traffic volume levels provided in Table 4 or 5 are recommended to be considered as signal guideline volumes for diamond interchanges.

CONCLUSIONS

The following conclusions were drawn from the data collected and field observations made within this study. They apply within the operational environment of one-way frontage roads.

1. Although each side of a diamond interchange is an intersection, a diamond interchange operates much differently than would two isolated intersections due to the close spacing.
2. Because diamond interchanges operate differently from isolated intersections, criteria for warranting diamond interchange signalization should be a separate MUTCD procedure from that for isolated intersections.
3. Diamond interchange models that uniquely combine the complex interactions of internal and external traffic appear to be the most representative approach on which to base diamond interchange guidelines for signalization.
4. There is a discriminating diamond interchange volume level beyond which traffic signal control is better than stop sign control in terms of the combined performance of queue and travel speed. The specific volume levels proposed for considering implementation of signalization at diamond interchanges are presented in Tables 4 and 5.

RECOMMENDATIONS

1. The guidelines presented in Tables 4 and 5 are recommended for implementation and testing to ascertain their acceptability for determining when and where installation of traffic signalizations is needed at diamond interchanges.
2. Separate signalization warrants for diamond interchanges are recommended. The guidelines provided in Tables 4 and 5 should be considered in the development of diamond interchange signal warrants in the MUTCD.
3. Further research is recommended to determine the exact number of hours during the average day that should meet the guideline volume levels for implementation purposes.

ACKNOWLEDGMENTS

The data in this paper are based on research sponsored by the Texas State Department of Highways and Public Transportation (TSDHPT) in cooperation with FHWA, U.S. Department of Transportation. Special thanks are extended to the following individuals and organizations for their help in providing assistance, cooperation, and technical comments for the study. Herman E. Haenel, Alvin R. Luedecke, and Blair G. Marsden of the Austin Headquarters of the TSDHPT; Stephen L. Cohen of FHWA; and traffic engineers and technicians in the cities of Abilene, Houston, and Fort Worth and in Districts 2, 8, and 12 of TSDHPT.

Finally, the authors would like to thank their fellow staff members at TTI. Contributions of the research technicians for data collection and reduction are sincerely appreciated. Particular appreciation is extended to graduate assistants Do Il Kim and Sai-Kwon Park, who provided consistent help throughout the project.

REFERENCES

1. Manual on Uniform Traffic Control Devices for Streets and Highways. FHWA, U.S. Department of Transportation, 1978.
2. Texas Manual on Uniform Traffic Control Devices for Streets and Highways. Texas State Department of Highways and Public Transportation, Austin, 1980.
3. C. Pinnell and D.C. Capelle. Operational Study of Signalized Diamond Interchanges. Bull. 324. HRB, National Research Council, Washington, D.C., 1962, pp. 38-72.
4. C.J. Messer, R.H. Whitson, and J.D. Carvell, Jr. Real-Time Frontage Road Progression Analysis and Control Strategy. In Transportation Research Record 503, TRB, National Research Council, Washington, D.C., 1974, pp. 1-12.
5. C.J. Messer, D.B. Fambro, and S.H. Richards. Optimization of Pretimed Signalized Diamond Interchanges. In Transportation Research Record 644, TRB, National Research Council, Washington, D.C., 1977, pp. 78-84.
6. P.K. Munjal et al. Design Manual for Traffic Signal Control of Diamond Interchange Complexes. Report TM-4601/007/03. FHWA, U.S. Department of Transportation, June 1972. NTIS: PB 217518.
7. Control Strategies for Signalized Diamond Interchange, Executive Summary. Report FHWA-TS-78-207. FHWA, U.S. Department of Transportation, April 1978.
8. Diamond Interchange Program--User's Manual. Report FHWA-IP-80-4. FHWA, U.S. Department of Transportation, Dec. 1980.
9. R. Jenkins and R. Wiles. Case Study of a Diamond Interchange. TexITE Newsletter, ITE, Washington, D.C., April 1982.
10. SAS User's Guide: Basics and Statistics. SAS Institute, Cary, N.C., 1982.

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

Optimal Timing Settings and Detector Lengths of Presence Mode Full-Actuated Control

FENG-BOR LIN

ABSTRACT

The operation of presence mode full-actuated signal control at individual intersections is governed primarily by the choice of detector length and the timing settings of vehicle interval and maximum green. The relationships between these control variables and the control efficiency vary with the flow pattern at an intersection. Based on the results of computer simulations, the optimal combinations of detector length, vehicle interval, and maximum green are identified for a wide range of flow conditions. The analyses performed in this study concern only intersections where vehicle approach speeds are less than 35 mph.

Full-actuated signals based on long loop presence detectors are being widely used for the regulation of traffic flows at individual intersections. This presence mode control, which is also referred to as loop-occupancy control, can rely on a variety of timing settings and detectors. Nevertheless, the typical operation of this mode of control is governed by three basic control variables: vehicle interval, maximum green, and detector length. Vehicle interval determines the longest duration in which detectors can be left unoccupied without prompting the termination of a green duration. Maximum green limits the maximum green duration allowable to a signal phase after a vehicle actuates a detector of a competing phase.

Some researchers have attempted to quantify the performance of the presence mode control under certain operating conditions, but so far the findings are inconclusive regarding the optimal use of presence mode control. for example, Cribbins and Meyer (1) examined the effects of detector length on the control efficiency under real-life conditions. They concluded that the longer the length of the presence detector on the major approach to an intersection, the longer the delay. The conditions under which various detector lengths were examined, however, are unknown.

To provide further insights into the effects of detector length, Tarnoff and Parsonson (2) used the NETSIM simulation model (3) to compare detector lengths of 30 to 90 ft. They found that the efficiency of the presence mode control increased as the detector length was shortened. But they also cautioned that the simulation results did not properly account for the possibility that a signal phase could be prematurely terminated because of the variations in queue discharge headways. Tarnoff and Parsonson's caution is not unwarranted. A recent study by Lin and Percy (4) has indicated that the risk of the premature phase termination is not negligible and can significantly affect the operating characteristics of the presence mode control.

Generally, current understanding of the performance characteristics of the presence mode control is piecemeal and mostly intuitive in nature. As a result, it is not clear how this mode of control can be used to achieve the highest possible control efficiency. To bridge this gap in the state of the art of signal control, this study was conducted to determine the relationships between the optimal use of the presence mode control and the flow patterns at individual intersections where vehicle approach speeds are less than 35 mph.

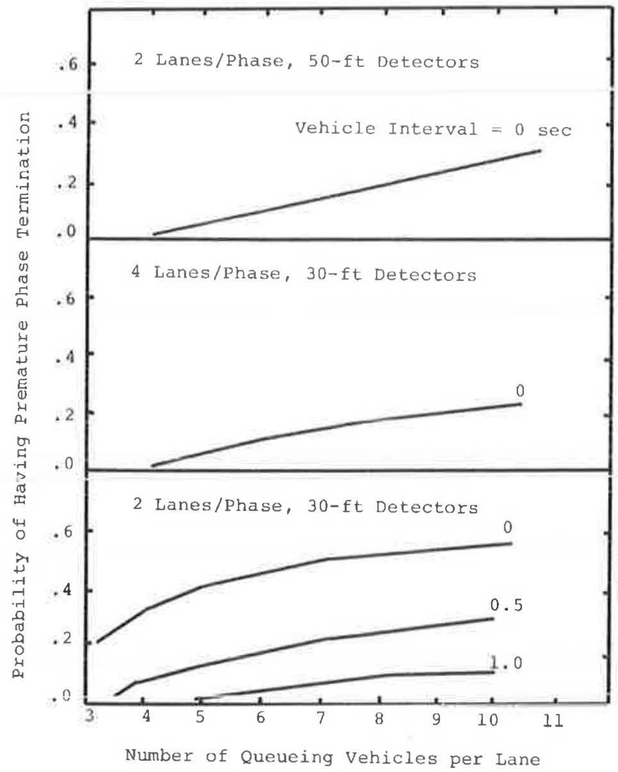


FIGURE 1 Probabilities that queuing vehicles face prematurely terminated green phases.

METHOD OF PERFORMANCE ANALYSES

It is generally impractical to conduct field studies to determine how the performance of the presence mode control would change under different operating conditions. A practical alternative is to use computer simulation for the performance analysis of such a signal control. But past efforts in developing simulation models largely overlooked the importance of a reasonably accurate representation of the interactions between queuing vehicles and presence detectors. Consequently, existing models may not be suitable for use as a tool to identify the optimal use of the presence mode control.

The NETSIM model (3), for example, assumes that every vehicle in a queue can extend a green duration until the queue dissipates completely. Field data collected in this study (5), however, indicate that this is not always the case. In fact, the probability of premature phase termination caused by the failure of queuing vehicles to extend green durations can be rather high, even when 50-ft detectors are used. Figure 1 presents a few examples of this phenomenon. The existence of prematurely terminated green durations can be expected to result in poor operation of the presence mode control. Ignoring it would certainly lead to underestimates of vehicle delays.

To avoid introducing systematic biases into the analysis of presence mode signal operations, a simulation model referred to as the RAPID model (5) was used in this study. This model is a microscopic simulation model capable of duplicating the dynamic and probabilistic interactions between queuing vehicles and presence detectors. This capability is indispensable because the performance of the presence mode control is dictated by such interactions. Furthermore, the model does not contain any assumption that would misrepresent the actual operation of the presence mode control.

To facilitate this study, the RAPID model was calibrated with field data collected at two intersections (5). These data concern the interactions between queuing vehicles and detectors 30 to 120 ft long. Based on this calibrated model, the optimal combinations of detector length, vehicle interval, and maximum green were determined for a variety of flow patterns. Vehicle delays were used as the measure of performance in search of such optimal combinations.

The average vehicle delay associated with a presence mode operation is a random variable. Its true value can only be estimated. Therefore, the optimal control referred to herein for a given flow pattern is in fact an approximate solution. The procedure used to search for such an optimal control was simple but tedious. Detector lengths of 30, 50, 65, 80, and 120 ft are evaluated separately first. For each detector length, the best combinations of vehicle interval and maximum green were identified for a number of flow patterns. The average delay produced by each combination of vehicle interval and maximum green was estimated on the basis of the outputs of at least four simulation runs. For a flow pattern with heavy lane flows, the average delay could vary substantially from one simulation run to another. In such a case additional simulation runs were performed to obtain a better estimate of the average delay.

The flow patterns examined in this study represent a number of combinations of flow rate per lane, distribution of traffic volume among lanes, and temporal variations in flow rate. The vehicles associated with these flow patterns included straight-through and right-turn movements with a negligible number of trucks and buses. These two directional movements represent two extreme flow conditions and were analyzed separately. Both two- and four-phase operations of the presence mode control were analyzed for each flow pattern. The rest-in-red feature was assumed to be in effect. Also, each signal phase contained up to four lanes. The ratio of the critical lane flow in one phase to that in another phase was varied from 1 to 2. The flow rate in a lane ranged from 50 to 100 percent of the critical lane flow of the same phase.

Each combination of flow pattern, detector length, vehicle interval, and maximum green was analyzed on the basis of a 1-hr operation of the signal control. In such an hourly operation, the flow rate in each lane was allowed to vary from time to time at 5-min intervals. A factor, referred to herein as the peaking factor (PF), was used to represent the degree of such temporal variations in the flow rate. This factor is defined as

$$PF = \text{Hourly volume}/(4 \times \text{peak 15-min volume}) \quad (1)$$

A peaking factor of 1.0 indicates a uniform flow rate. A lower peaking factor implies that there is a higher concentration of traffic in a short period of time. The signal operation for each flow pattern was analyzed, respectively, at peaking factors of 1.0, 0.85, and 0.7. At any flow rate, the arrivals of the vehicles were assumed to be random.

PERFORMANCE CHARACTERISTICS

The choice of detector length, vehicle interval, and maximum green can affect the risk of the premature phase termination. It can also affect the speed profile of a vehicle before and after the actuation of a detector and the degree of easiness or difficulty for vehicles not in a queue to extend a green phase. The resulting relationships between the control efficiency and the control variables are complex.

Among the three control variables, maximum green plays a relatively simple and easily identifiable role in shaping the operation of the presence mode control. The maximum green chosen for a specific signal operation is dormant until the arriving vehicles are able to extend a green phase continuously. The potential impact of this signal control variable is shown in Figure 2.

One feature revealed in this figure is that the delays are insensitive to maximum green when the flows are relatively low [e.g., 400 vehicles per hour (vph) per lane]. For a flow pattern with heavier flows, short maximum greens become undesirable. In such a case the average delay begins to increase rapidly when the maximum greens fall below a certain level. Long maximum greens, however, may not have a significant adverse impact on the average delay.

Figure 2 also shows that optimal maximum green can vary with the peaking factor. The general trend as revealed by the simulation data is that the optimal maximum green decreases when the peaking factor increases. This is not an unexpected result. A flow pattern with a strong peaking characteristic (i.e., small peaking factor) implies a high concentration of traffic volume in a short period of time. For a given hourly flow rate, the smaller the peaking factor, the heavier the traffic becomes in such a period and the longer the maximum green should be.

The effects of detector length and vehicle interval on control efficiency are much more difficult to generalize. Nevertheless, the operation of the presence mode control is governed primarily by the sum of the dwell time of a vehicle in a detection area and the vehicle interval provided. This sum can be referred to as the effective vehicle interval

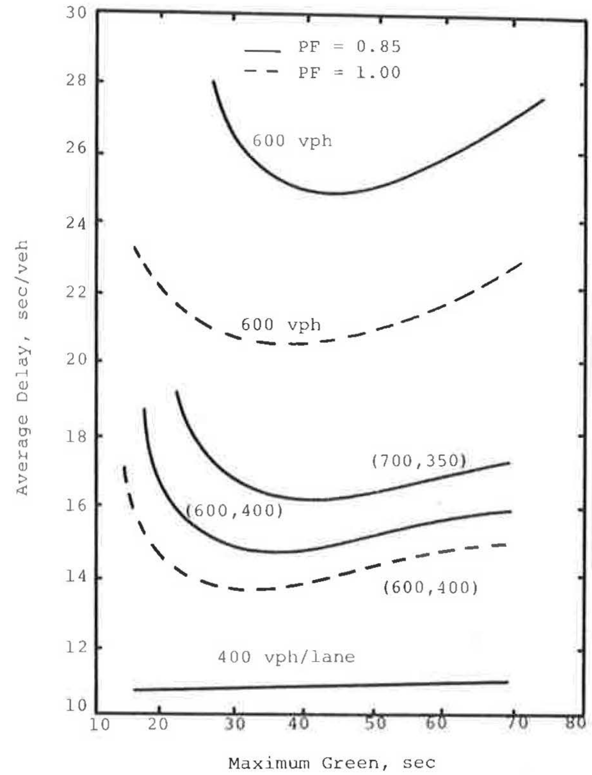


FIGURE 2 Average delay as a function of maximum green and lane flow for two-phase operations.

faced by a vehicle. The dwell time is a function of detector length and can vary from one vehicle to another. Consequently, the effective vehicle intervals faced by the arriving vehicles also vary.

If such effective vehicle intervals are short, both queuing vehicles and vehicles not in a queue may have great difficulties extending a green phase. On the other hand, long effective vehicle intervals may allow vehicles separated by long headways to extend a green phase. In either case, control efficiency can be expected to be poor.

After a green phase begins, the queue in a lane will grow and decay at the same time. Eventually such a queue will dissipate. For the vehicles in a queue, the premature termination of a green phase should be prevented. Otherwise the queue length may grow from one cycle to another, thus inducing excessive delays. Under highly variable flow conditions, the premature phase termination can be effectively prevented if long detectors (e.g., 80 ft) are used. Once the risk of the premature phase termination is negligibly small because of the use of long detectors (e.g., 80 ft or longer) or because of the presence of light traffic flows, then longer vehicle intervals may lead to increased delays. These characteristics of the presence mode control are clearly revealed in Figure 3.

After a queue dissipates from a detection area, it may be desirable to allow some vehicles that are following behind to extend the green phase. The vehicles not in the queue generally have longer headways than the queuing vehicles in the same lane. Furthermore, they are faced with effective vehicle intervals that are shorter than those encountered by the queuing vehicles. Consequently, such vehicles can be expected to have substantial difficulties in extending a green phase. An approximate analysis given in the following paragraphs underscores this phenomenon.

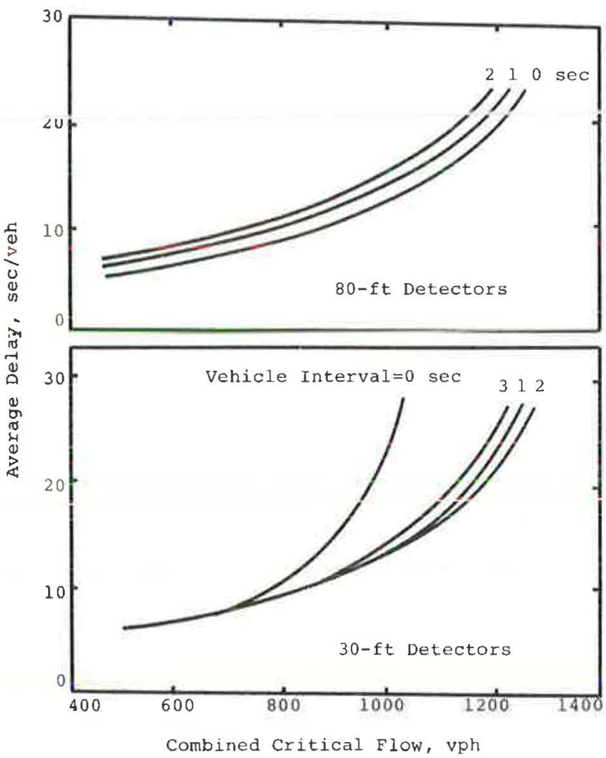


FIGURE 3 Variations in delays with combined critical flow and vehicle interval (two phases, four lanes per phase, PF = 0.85).

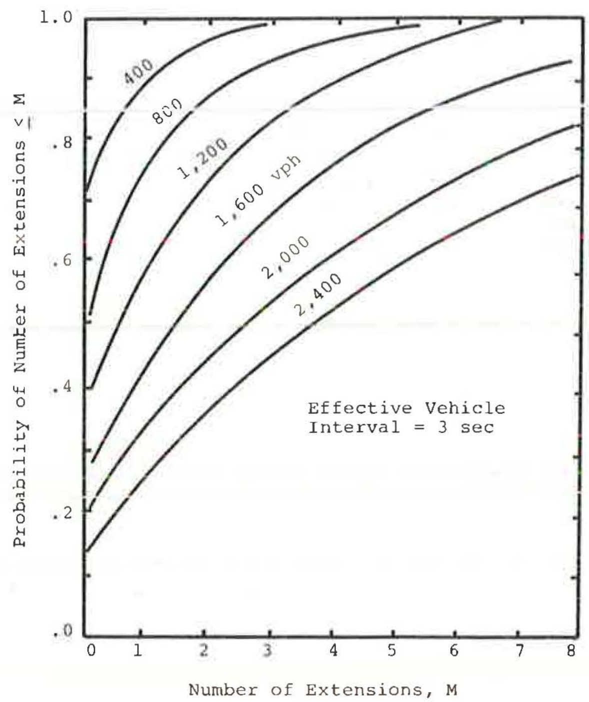


FIGURE 4 Probabilities that M or fewer vehicles not in a queue will be able to extend a green phase.

Consider a green phase that is associated with more than one lane and assume that vehicles that cannot join a queue arrive randomly at the upstream end of the detector in a lane. With these random arrivals, it can be shown (6) that the arrival headways of the combined flow can be represented by the following probability density function:

$$F(h \geq t) = \exp[-(\lambda_1 + \lambda_2 + \ldots + \lambda_i \ldots)t] = e^{-\lambda t} \quad (2)$$

where

$F(h \geq t)$ = probability that a headway h is greater than or equal to t,
λ_i = flow rate in lane i, and
λ = combined flow rate.

As an approximation, let the effective vehicle interval faced by each of such vehicles be the same. Denote this effective vehicle interval as U. Then, for a vehicle in the combined flow to extend a green phase, its headway should not exceed U. The probability that a vehicle will be able to extend the green becomes $1 - e^{-\lambda U}$. The corresponding probability Y that exactly M vehicles will be able to extend the green in succession is

$$Y = (1 - e^{-\lambda U})^M e^{-\lambda U} \quad (3)$$

Based on this equation, the probability that M or fewer vehicles will be able to extend a green phase can be estimated for various combinations of λ and U. Figure 4 shows that, with an effective vehicle interval of 3 sec, the median number of vehicles that can extend a green phase in succession is only about 2.5 when the combined flow is 2,000 vph. If the effective vehicle interval is increased to 4 sec, it can be shown that the corresponding median value is still only five vehicles. This is an average of 1.25 vehicles per lane if the combined flow is distributed among four lanes. Therefore, unless the effective vehicle intervals faced by the arriving vehicles are longer than 4 sec and the combined flow is extremely heavy, the presence mode control will rarely allow a vehicle not in a queue to extend a green phase.

Allowing vehicles not in a queue to extend a green phase is desirable only when the combined critical flow of a traffic pattern is heavy. With a combined critical flow exceeding 1,200 vph, for example, 1-sec vehicle intervals tend to produce more efficient operations than 0-sec vehicle intervals when 65-ft detectors are used. No field observations have been made on vehicle movements over 65-ft detectors. Nevertheless, the probability of the premature phase termination associated with the use of such detectors can be expected to be negligibly small. This implies that the 1-sec vehicle intervals needed to minimize delays are primarily to allow some vehicles not in a queue to extend a green phase.

OPTIMAL UTILIZATION

Optimal Maximum Green

Maximum green is usually set between 30 and 60 sec (7). Current practices in selecting the maximum green appear to be arbitrary. For the purpose of preventing a green phase from becoming unreasonably long to waiting drivers, the maximum green may be set in accordance with a tolerable waiting time. How long a waiting time is tolerable is, of course, subject to intuitive judgment.

To maintain high control efficiency under varying flow conditions, it has also been suggested (7) that the maximum green be selected to correspond to the desired cycle length and split at an intersection. Following this suggestion, the optimal pretimed cycle length and green durations for a flow pattern being considered can be determined first. The computed green durations are then multiplied by a factor ranging between 1.25 and 1.50 to obtain the maximum greens (7). This approach is logical, but the basis for choosing a value between 1.25 and 1.50 as the multiplication factor is not clear.

The simulation results obtained in this study indicate that the optimal maximum green of a phase can be related to the corresponding optimal pretimed green and the peaking factor. Figure 5 shows such relationships. The optimal pretimed green, denoted

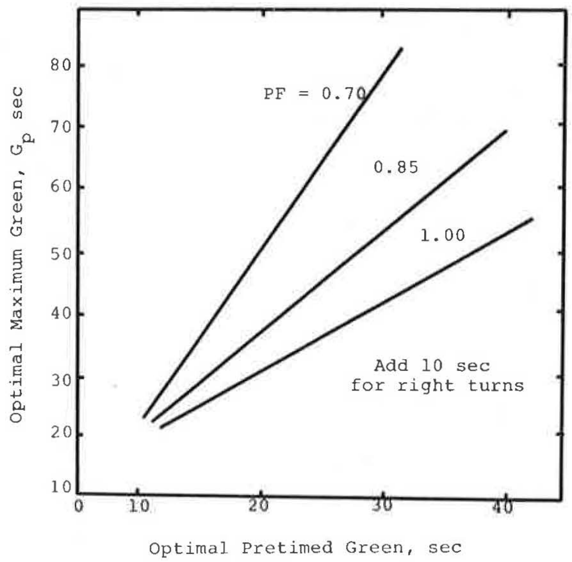


FIGURE 5 Recommended maximum greens.

as G_p in the figure, is determined from the following optimal pretimed cycle length (8):

$$C_o = (1.5\,L + 5) \bigg/ \left(1 - \sum_{j=1}^{N} Z_i\right) \qquad (4)$$

where

C_o = optimal pretimed cycle length (sec);
L = loss time per cycle, taken as 5 sec per phase;
Z_i = ratio of critical lane volume of phase i to saturation flow of 1,800 vph; and
N = number of signal phases.

Given C_o, the available green time is allocated to each phase in proportion to the critical lane volume; that is,

$$G_p = \left(C_o - \sum_{j=1}^{N} Y_j\right) \bigg/ \left(Q_{ci} / \sum_{j=1}^{N} Q_{cj}\right) \qquad (5)$$

where

G_p = pretimed green duration of phase i,
Y_j = clearance interval of phase j, and
Q_{ci},Q_{cj} = critical lane volumes of phase i and phase j, respectively.

A few observations can be made from Figure 5. First, with a peaking factor of 1.0, the optimal maximum greens are about 10 sec longer than the corresponding optimal pretimed greens. Second, when the peaking factor decreases to 0.85, the optimal maximum greens are about 80 percent longer than the pretimed greens. Finally, the optimal maximum greens for a peaking factor of 0.7 are about 2.5 times the optimal pretimed greens. The optimal maximum greens for flow patterns with only right-turn flows are longer than those for straight-through flows. The difference is about 10 sec.

The consequences of using maximum greens that deviate from the values given in Figure 5 are shown in Figure 6. Each curve of this figure represents the average delays for a given flow pattern when maximum green is varied. It is obvious from this figure that the use of maximum greens 10 sec shorter than the values given in Figure 5 should be avoided. On the other hand, maximum greens 20 sec longer than such values may increase average delays only slightly.

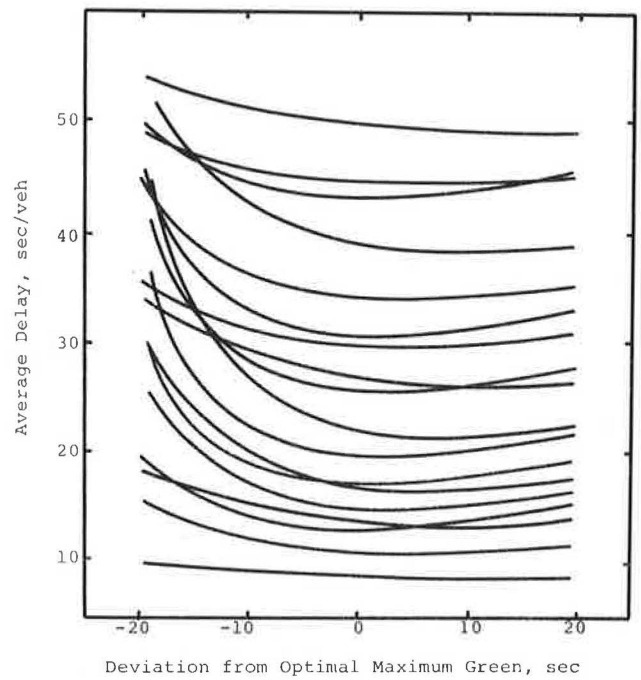


FIGURE 6 Variations in average delays as a function of maximum green.

Because the traffic volume at an intersection varies from time to time, the optimal maximum green of a signal phase should be determined on the basis of the peak-hour flow pattern when only one maximum green per phase is allowed. If two settings of maximum green are allowed, one setting should be based on the peak-hour flow patterns and the other based on a pattern with moderate flow rates. The maximum green should be limited by drivers' tolerance to waiting.

Optimal Vehicle Interval

For detectors at least 80 ft in length, 0-sec vehicle intervals can be expected to produce the most efficient signal operations. When shorter detectors are used, the optimal vehicle intervals depend primarily on the combined critical flow of the traffic pattern. A heavier combined critical flow generally requires a longer vehicle interval in order to minimize delays.

When 1-sec vehicle intervals are chosen over 2-sec vehicle intervals for 30-ft detectors, Figure 7 shows that the average delays may be reduced by up to 2 sec per vehicle for straight-through flows and up to 4 sec per vehicle for right-turn flows if the combined critical flow is less than 900 vph. Under heavier flow conditions, 2-sec vehicle intervals become much more desirable than 1-sec vehicle intervals. For straight-through flows with a combined critical flow of more than 1,000 vph, there is an increasing need to use 3-sec vehicle intervals. For right-turn flows it becomes advantageous to use 3-sec vehicle intervals only when the combined critical flow approaches 1,400 vph. The use of vehicle intervals longer than 3 sec, however, would reduce control efficiency.

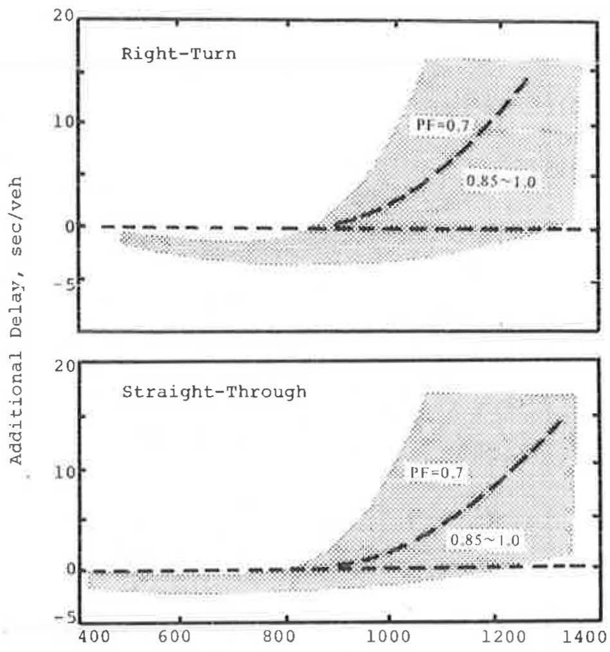


FIGURE 7 Additional delays caused by the choice of 1-sec vehicle intervals over 2-sec vehicle intervals (30-ft detectors).

For 50-ft detectors serving straight-through flows, 1-sec vehicle intervals are always better than 2-sec vehicle intervals when the combined critical flow is less than 1,000 vph. Above this flow level the relative efficiencies of 1- and 2-sec vehicle intervals depend on the specific flow pattern at an intersection. Generally, 2-sec vehicle intervals can become slightly better when the peaking factor approaches 0.7, whereas 1-sec vehicle intervals are preferred when the peaking factor is greater than 0.85. The differences in the resulting delays, however, are less than 1.5 sec per vehicle and thus can be ignored. When the combined critical flow is less than 800 vph, 0-sec vehicle intervals are preferred to 1-sec vehicle intervals. Under heavier flow conditions, it becomes important to use 1-sec vehicle intervals.

To serve right-turn flows, there is no advantage of using vehicle intervals longer than 1 sec for 50-ft detectors. For such directional flows, 0-sec vehicle intervals produce more efficient control than 1-sec vehicle intervals when the combined critical flow is less than 900 vph. Once the combined critical flow exceeds 900 vph, there is an increasing need to use 1-sec vehicle intervals.

The optimal vehicle intervals for 65-ft detectors are in the range of 0 to 1 sec. The use of longer vehicle intervals can be expected to induce additional delays. For 65-ft detectors used to serve straight-through flows, 0-sec vehicle intervals are preferred to 1-sec vehicle intervals when the combined critical flow is less than 1,000 vph. Under heavier flow conditions, 1-sec vehicle intervals should be used. For right-turn flows, the use of 0-sec vehicle intervals is generally desirable when the combined critical flow is less than 1,100 vph. Above this level of combined critical flow there is an increasing need to use 1-sec vehicle intervals.

Based on these findings, a set of vehicle intervals is determined and recommended for timing design applications. These recommended vehicle intervals are shown in Figure 8. The shaded areas in this figure represent various ranges of vehicle intervals in which the control efficiency is not likely to vary significantly. Nevertheless, it is desirable to use the upper bounds of such ranges to choose a vehicle interval when the peaking factor of a flow pattern approaches 0.7. The lower bounds are to be used when the peaking factor is between 0.85 and 1.0. As in the case of selecting a maximum green, the timing design can be based on the peak-hour flow pattern expected at an intersection.

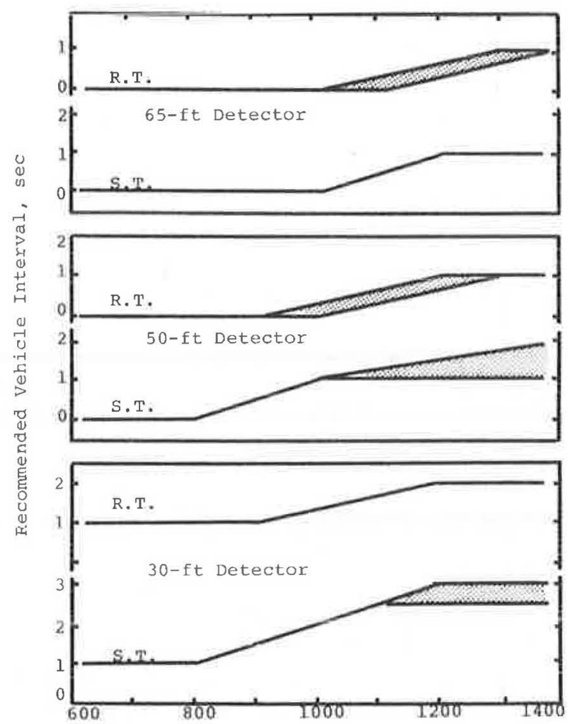


FIGURE 8 Recommended vehicle intervals.

For detector lengths not shown in Figure 8, their optimal vehicle intervals can be estimated through interpolations.

Optimal Detector Length

It has been suggested (9) in the past that required detector lengths be determined from the following equation:

$$D = 1.47\ V(U - E) - L \quad (6)$$

where

D = detector length (ft),
V = design approach speed (mph),
U = desired effective vehicle interval (sec),
E = vehicle interval (sec), and
L = design vehicle length (ft).

The primary concern of this equation is to give vehicles not in a queue a reasonable chance to extend a green phase. To serve this purpose, Equation 6 in fact equates the sum of the dwell time $(D + L)/(1.47V)$ of such vehicles and the vehicle interval E to a desired effective vehicle interval U. An effective vehicle interval of 3 sec is usually considered to be adequate.

Equation 6 is convenient to use, but the detector lengths determined from it are unlikely to be the most desirable. A major reason for this is that the effective vehicle interval U that should be used in the equation has not been clearly specified. For a detector of 30 to 65 ft, it has been shown (Figure 8) that the vehicle interval needed to minimize delays increases with the combined critical flow of a flow pattern. This implies that the effective vehicle interval U should be related to traffic volume. Furthermore, the operation of the presence mode control is governed primarily by the interactions between queuing vehicles and detectors. Therefore, the use of Equation 6 may lead to good choices of detector length for some flow patterns and poor choices for others.

When compared with the other detector lengths examined in this study, 80-ft detectors with 0-sec vehicle intervals were found to be able to produce either better or at least equally efficient signal operations. Figure 9 provides an insight into the relative efficiencies of the various detector lengths. Each delay curve shown in the figure is associated with a flow pattern under either a two- or four-phase signal control. It can be seen from the figure that the control efficiencies produced by 65- and 80-ft detectors are comparable over a wide range of flow conditions.

Detectors shorter than 80 ft are frequently used because of budget constraints. Before such detectors are used, their potential impact on control efficiency should be evaluated. Figure 10 shows an example of the additional delays that may result from the use of detectors shorter than 80 ft. The information contained in a figure such as this can be used to assist in the choice of detector lengths. For example, the data in Figure 10 indicate that, when the combined critical flow is less than 1,000 vph, the average delays caused by the use of 65-ft detectors are less than 1.5 sec per vehicle longer than those produced by the use of 80-ft detectors. Therefore, if this magnitude of the added delays is deemed to be insignificant, then 65-ft detectors may be employed.

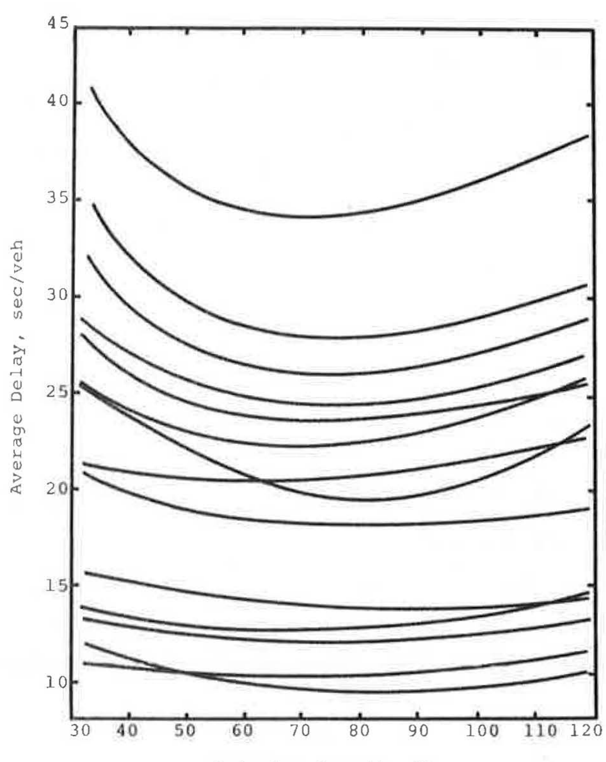


FIGURE 9 Variations in average delays with detector length.

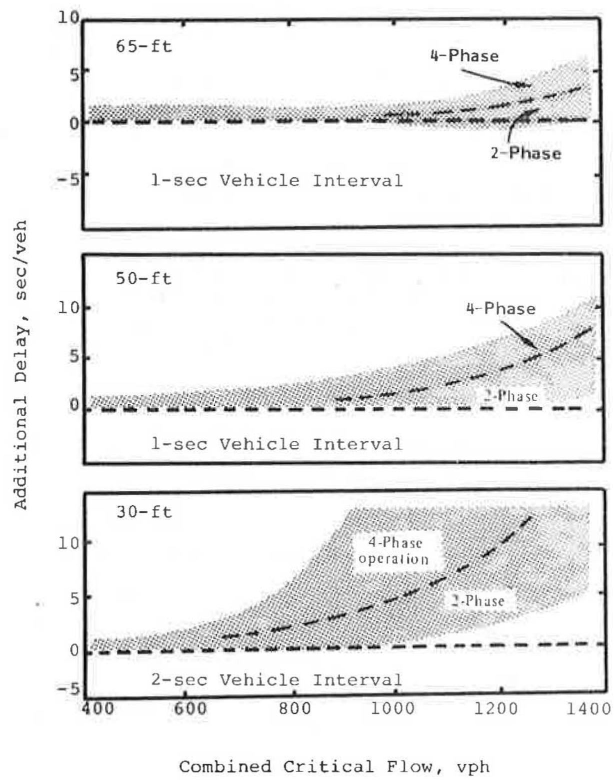


FIGURE 10 Additional delays caused by the use of detectors shorter than 80 ft (straight-through flows).

Short detectors may be used in place of 80-ft detectors as long as their use is not likely to incur undue additional delays. Figures 11 and 12 show such acceptable detector lengths at two levels of allowable added delays for straight-through and right-turn flows, respectively. These figures indicate that, to maintain a specified level of control efficiency, detector length should increase with combined critical flow. They also reveal that, if an added delay of 5 sec per vehicle is acceptable, 30-ft detectors can be used for flow patterns with combined critical flows of up to about 800 vph. Four-phase operations require longer detector lengths than two-phase operations because they incur longer delays, and such delays are more sensitive to the choice of detector length.

The detector lengths determined from Figures 11 and 12 should be used in conjunction with the

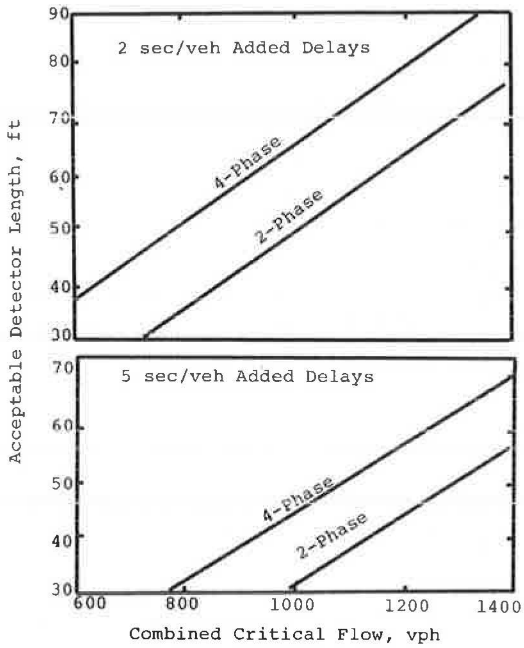


FIGURE 11 Acceptable detector lengths and allowable added delays (straight-through flows).

optimal vehicle intervals for respective levels of combined critical flow. Otherwise, the resulting added delays may be significantly longer than the acceptable values indicated in these figures.

CONCLUSIONS

Generally, optimal maximum greens for the presence mode control are longer than the green durations required to achieve optimal pretimed control. Flow patterns with higher degrees of concentration of traffic in short periods of time need longer optimal maximum greens. The optimal maximum greens for hourly flow patterns with a peaking factor of 1.0 are about 10 sec longer than the corresponding optimal pretimed greens. With a peaking factor of 0.85, the optimal maximum greens are approximately 80 percent longer than the corresponding optimal pretimed greens.

Optimal vehicle intervals are a function of detector length and flow rate. For detectors 30 ft long, the use of 2-sec vehicle intervals can lead to the best signal performance over a wide range of operating conditions. For 50-ft detectors, 1-sec vehicle intervals are desirable under a variety of flow conditions. When detectors 80 ft or longer are used, 0-sec vehicle intervals can minimize delays. The use of vehicle intervals longer than 0 sec for such detector lengths is not desirable unless the combined critical flow at an intersection exceeds 1,400 vph.

Detectors 80 ft long can consistently produce the best signal performance. For a combined critical flow of less than 1,100 vph at an intersection, however, 65-ft detectors can produce comparable performance. For a combined critical flow of less than 900 vph, the use of 50-ft detectors in place of 80-ft detectors would only incur an added delay of up to 2 sec per vehicle. And, for a combined critical flow of less than 600 vph, 30-ft detectors may also be used to replace 80-ft detectors without incurring undue excess delays. For a specific signal control problem that involves the use of presence detectors, it is recommended that Figures 5, 8, 11, and 12 be consulted to determine an efficient combination of detector length, vehicle interval, and maximum green. For a signal phase with various directional flows, the critical lane flow of a representative peak-hour flow pattern may be used as a basis for determining such a combination.

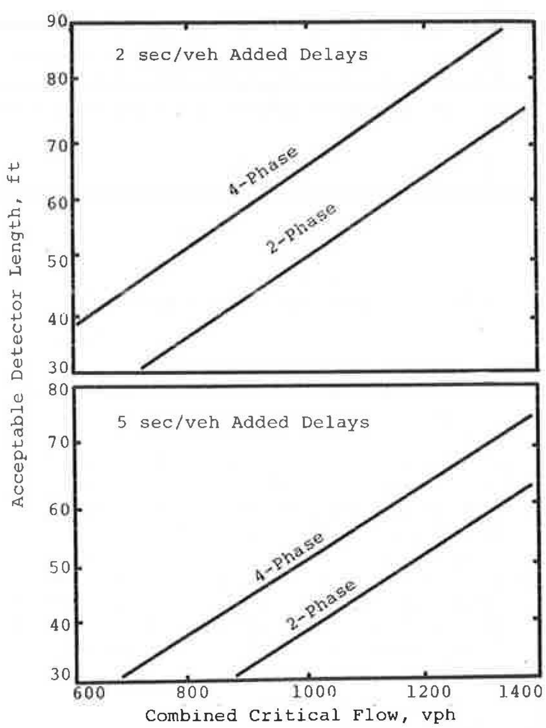


FIGURE 12 Acceptable detector lengths and allowable added delays (right-turn flows).

ACKNOWLEDGMENTS

The author wishes to thank Steven Shen, a former graduate student at Clarkson University, for his assistance in collecting field data and in performing simulation analyses. Thanks are also extended to Carl Davey of the Civil and Environmental Engineering Department of Clarkson University for his instrumentation of electronic equipment used in the data collection. This study was sponsored by the Office of University Research, U.S. Department of Transportation.

REFERENCES

1. P.E. Cribbins and C.A. Meyer. Choosing Detectors and Placing Them Right. Traffic Engineering, Vol. 45, No. 1, Jan. 1975, pp. 15-18.
2. P.J. Tarnoff and P.S. Parsonson. Selecting Traffic Signal Control at Individual Intersections. NCHRP Report 233. TRB, National Research Council, Washington, D.C., June 1981, 133 pp.
3. Traffic Network Analysis with NETSIM--A User Guide. Report FHWA-IP-80-3. FHWA, U.S. Department of Transportation, Jan. 1980.
4. F.-B. Lin and M.C. Percy. Vehicle-Detector Interactions and Analysis of Traffic-Actuated Signal Controls. In Transportation Research Record 971, TRB, National Research Council, Washington, D.C., 1984, pp. 112-120.
5. F.-B. Lin. Optimal Timing Settings and Detector Lengths of Presence Mode Full-Actuated Control.

Draft Final Report. Office of University Research, U.S. Department of Transportation, Dec. 1984.
6. D.L. Gerlough and F.C. Barnes. Poisson and Other Distributions in Traffic. ENO Foundation for Transportation, Saugatuck, Conn., 1971.
7. J.H. Kell and I.J. Fullerton. Manual of Traffic Signal Design. Prentice-Hall, Englewood Cliffs, N.J., 1982.
8. F.V. Webster. Traffic Signal Settings. Road Research Technical Paper 39. Road Research Laboratory, London, England, 1958, pp. 1-44.
9. ITE Southern Section Technical Committee 18. Large-Area Detection at Intersection Approaches. Traffic Engineering, June 1976, pp. 28-36.

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

Multiway Stop Sign Removal Procedures

CLAUDE M. LIGON, EVERETT C. CARTER, and HUGH W. McGEE

ABSTRACT

In recent years local jurisdictions have successfully converted unwarranted multiway stop-controlled intersections to less restrictive forms of control. However, there is wide variation in the approaches used and factors considered in the conversion decision. Therefore, FHWA initiated a national study of the processes, with two primary objectives: (a) to develop and test procedures to convert multiway stop-sign-controlled intersections to two-way stop-sign-controlled intersections, and (b) to document the safety effects of converting multiway stop controls to two-way controls. In this paper the study is summarized and the results are presented in the form of recommended conversion procedures. Thirty separate geographically distributed jurisdictions were visited and information and data regarding the various conversion experiences were collected. Data from more than 170 separate intersections were studied by the research team in arriving at the conclusions and recommended procedures in this paper. Laboratory driver preference studies were conducted to determine the most suitable warning and information signs. In addition to local government officials, several consultants as well as professionals in quasi-public agencies were interviewed and their experiences and knowledge of the conversion process were incorporated, where appropriate. The emphasis of the study has been on the safety aspects of the conversion process.

Within the past few decades there has been an increase in the use of multiway stop signs as the traffic control scheme at many intersections. Many elected officials believe that multiway stop signs are a panacea for intersection safety problems because they promote speed control, accident reduction, and pedestrian safety. Even though the Manual on Uniform Traffic Control Devices (MUTCD) (1) has warrants for the application of multiway stop control, in some cases the "political" warrant is the only one that is met. Multiway stop signs should ordinarily be used only where the intersecting road volumes are approximately equal. The MUTCD states that a stop sign should not be used for speed control.

Research has indicated that stop signs installed to control speed do not result in speed reduction (2-5). Also, studies have indicated that stop signs do not always result in increased safety (6).

Unwarranted stop signs increase stops, cause delays, and increase fuel consumption and pollutants. Further, installation of unwarranted traffic control devices breeds disrespect for such devices and can result in potentially dangerous behavior. For these reasons, it is desirable to remove unwarranted and unneeded stop signs that hinder traffic flow rather than aid it. Concern for the environment and for fuel conservation has led to a different attitude toward traffic control.

For several decades traffic engineering changes have, almost without exception, involved installing more positive or rigid control; for example, going from no control to two-way stop control or two-way to four-way stop control. Traffic engineers as well as the general public are conditioned to increasing degrees of control. Local jurisdictions are beginning to realize the mistakes of the past and understand that there are air pollution, delay, and energy impacts that result from excessive use of multiway stops.

A recent study (7) indicates that pedestrian and vehicle accidents may increase when certain traffic volumes, intersection configurations, and approach speeds are combined.

PROJECT OBJECTIVE AND SCOPE

This study was undertaken with two primary objectives in mind:

1. To develop and test procedures to convert multiway stop-sign-controlled intersections to two-way stop-sign-controlled intersections, and
2. To document the safety effects of converting multiway stop controls to two-way controls.

The general approach was to visit at least 30 political jurisdictions that had multiway stop sign conversion experience. From their collective past experiences, and from methods that appeared to be reasonable, a recommended procedure was developed to convert multiway stop intersections to lesser forms of control.

DATA COLLECTION

Each political jurisdiction selected for a site visit designated those intersections that had been converted from multiway stop sign control to lesser forms of control. Data on the sites and on the number of intersections so identified are given in Table 1.

Data on the average daily traffic (ADT) and the posted speeds of the converted intersections studied are given in Tables 2 and 3, respectively. The fact that more than one-half of the converted intersections had ADTs of less than 1,500 vehicles per day and posted speeds of 25 mph or less suggests that most conversions identified in this study had been accomplished at residential intersections. This is often where complaints of speeding are most common and the "political" warrant for multiway stop sign installation is exercised. This situation often creates a climate for wholesale stop sign removals when subdivisions are annexed by a larger urban area because subdivisions often use stop signs as speed control devices.

Special signing for conversions was found to run the gamut in sizes and wording. Figures 1-4 contain some examples of signs used by various jurisdictions to assist in the conversion process.

ACCIDENT ANALYSIS

Because of the concern for the safety effects of converting multiway stop controls to two-way stop controls, an analysis of changes in accidents before and after conversions was conducted by using data for 172 intersections representing 33 jurisdictions

TABLE 1 Political Entities Contributing to Multiway Stop Sign Study

Political Entity	County/Parish	Population (000s)	No. of Converted Intersections Studied
FHWA Region 1			
Manchester, Conn.	Hartford	50	27
Colonie, N.Y.	Albany	78	3
Niskayuna, N.Y.	Schenectady	18	3
Troy, N.Y.	Rensselaer	56	8
FHWA Region 4			
Palm Beach County, Fla.	–	–	3
West Palm Beach, Fla.	Palm Beach	63	4
FHWA Region 5			
Berkley, Mich.	Oakland	20	5
Beverly Hills, Mich.	Oakland	12	2
Madison Heights, Mich.	Oakland	35	5
Trenton, Mich.	Wayne	25	5
Dayton, Ohio	Montgomery	200	7
FHWA Region 6			
Baton Rouge, La.	East Baton Rouge	250	2
Bossier City, La.	Bossier	55	20
Lafayette, La.	Lafayette	82	2
Oklahoma City, Okla.	Oklahoma	450	2
Arlington, Tex.[a]	Tarrant	160	6
Bellaire, Tex.	Harris	15	3
Houston, Tex.	Harris	1,500	3
Pasadena, Tex.	Harris	120	3
Seabrook, Tex.	Harris	5	2
Sugarland, Tex.	Fort Bend	9	15
Taylor Lake Village, Tex.	Harris	4	4
West University Place, Tex.	Harris	12	2
FHWA Region 7			
Olathe, Kans.	Johnson	39	4
Overland Park, Kans.	Johnson	82	4
Kansas City, Mo.	Jackson	448	5
FHWA Region 8			
Butte-Silverbow, Mont.[a]	Silverbow	37	9
FHWA Region 9			
Inglewood, Calif.	Los Angeles	90	4
Pamona, Calif.	Los Angeles	100	2[b]
Riverside, Calif.	Riverside	171	2
Riverside County, Calif.	–	–	2
San Bernardino, Calif.	San Bernardino	130	6
San Bernardino County, Calif.	–	–	1[b]

[a]Not included in site visits.
[b]Accident data not available.

TABLE 2 ADT of Converted Intersections

Total Intersection ADT Range	No. of Converted Intersections Studied	Percentage of Total Intersections
< 1,500	98	57
1,500–3,000	32	19
> 3,000	42	24
Total	172	100

TABLE 3 Posted Speeds of Converted Intersections

Speed (mph)	No. of Intersections Posted	Percentage of Total Intersections
20	6	3
25	101	59
30	49	29
35	6	3
40	9	5
50	1	1
Total	172	100

FIGURE 1 Example of advance motorist warning (Lafayette, Louisiana).

FIGURE 2 Supplementary notice sign (Baton Rouge, Louisiana).

FIGURE 3 Supplementary sign after conversion (Kansas City, Missouri). Note supplementary sign on post on opposite corner.

FIGURE 4 Pavement markings and STOP AHEAD sign used to emphasize presence of remaining stop sign after conversions (San Bernardino County, California).

in 12 states. Because of data limitations, all accident types were grouped and the primary data element was the number of accidents before and after the conversions. Accident rates could not be used for this analysis because for many locations the volume data for both periods were not available. Based on information provided by the various agencies, it was reasonable to assume nearly equal volumes before and after conversion. Accident summary statistics are given in Table 4. Results of an analysis, using the Statistical Program for Social

TABLE 4 Accident Summary Statistics

		Supplementary Sign	
	Total	Yes	No
No. of accidents before	88	77	11
No. of accidents after	144	101	43
Total (all intersections)	232	178	54
No. of intersections with increased accidents	28	13	15
No. of intersections with decreased accidents	16	12	4
No. of intersections with no change	128	32	96
Total	172	57	115

Sciences (SPSS) computer package, indicated the following:

1. There was significant increase in the number of accidents (based on the Poisson distribution test) after the conversion. Although the aggregate effect was a significant increase in accidents, only 16 percent of the 172 sites experienced an increase and 9 percent experienced a decrease. This finding indicates that there might be certain geometric or operating characteristics that determine whether an increase in accidents will occur.
2. The percentage increase in accidents was significantly higher where there were no supplementary signs, based on a chi-square test.
3. Seventy-four percent of the intersections (128 of 172) had no change in the number of accidents.

At those sites where the accidents increased, another accident analysis was performed to determine how soon the accidents occurred after the conversion took place. It was expected that there might be an unusually high incidence of accidents immediately after the conversion with a return to a normal situation after the motorist had become fully aware that the intersection was a two-way stop control. This analysis considered the number of accidents that occurred for each of 12 months before and after the conversion for five sites combined. It appears that if accidents do increase, there is a concentration of accidents occurring within the first month. The remainder of the accidents occurred throughout the balance of the year, with the fluctuations expected of normal accident occurrence.

With regard to the issue of whether or not accident frequency changes as a result of the conversion, no generalized conclusions can be drawn. In aggregate, there was a significantly higher number of accidents, and more intersections increased in accidents rather than decreased. No positive relationships could be determined between any operational or geometric factors and accident change for the limited data available. However, it is noted that at none of the locations that experienced a high increase in accidents was there low traffic volume (less than 1,500 ADT for the total intersection).

There is evidence that the first month immediately after the conversion is the most critical period for accident increase. Motorists who had traveled through the intersection frequently when under a multiway control expect the opposing traffic to stop. Even after the conversion, this expectation can linger.

The use of supplemental signs is intended to overcome this expectation. By advising motorists that in the future the conversion will take place at a certain time, and after the conversion has taken place warning motorists on the stop-controlled approaches that the other approaches do not require a stop, it is hoped that motorists will quickly adapt to the new system.

In regard to the effect of supplementary signs, the results of the analysis were conflicting. On the one hand, where signs were used, there was a greater percentage of sites where accidents decreased, and, overall, there was a smaller percentage increase in accidents compared with sites without signs. However, what cannot be ascertained is what further increase in accidents might have occurred if the signs had not been used.

EVALUATION OF SUPPLEMENTARY SIGNS

To test warning and information signs and advance notice signs, several alternative warning signs were considered. Based on the data collected from the 172 intersections where multiway stop signs had been removed (on suggestions by state, county, and municipal agencies), on reviews of the literature, and on discussions with members of the research team, seven different sign messages were formulated. These were tested with about 30 participants at the University of Maryland. As a result of this preliminary preference test, four signs were fabricated by the Baltimore Department of Transit and Traffic. Once these were fabricated, slides of these signs, together with slides taken at actual field locations, formed the basis for a laboratory experiment to test both the meaning and the motorist's preferences from among 11 sign message alternatives (see Figure 5).

The actual laboratory experiment was developed in two parts. Part I tested sign meaning. It consisted of slides of a four-way stop intersection (before) and the same intersection as a two-way stop (after), followed by slides of each of the alternative sign messages for warning and information.

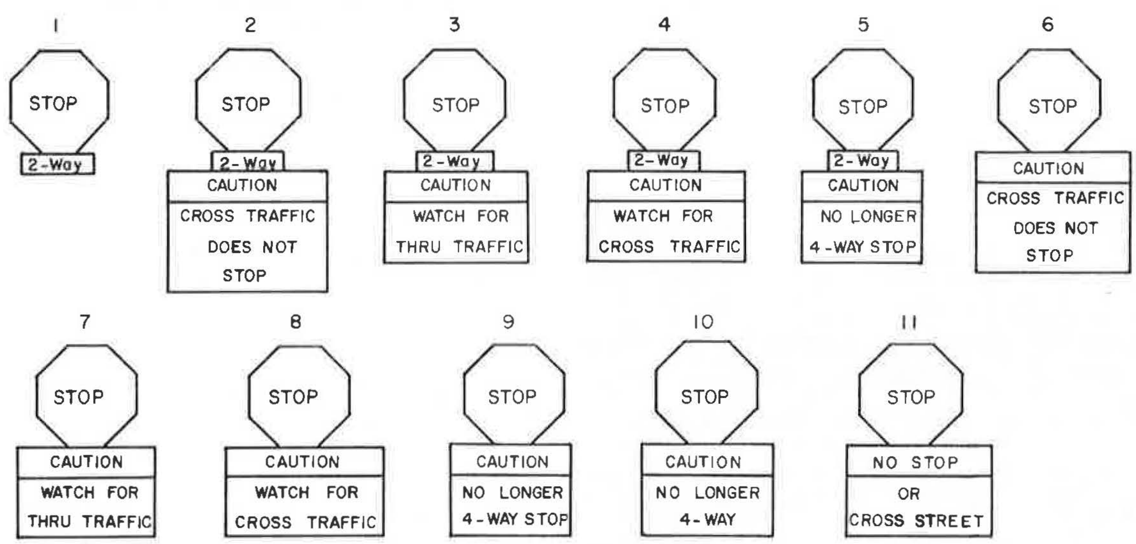


FIGURE 5 Sign messages selected for laboratory test evaluation.

A series of four answers were developed for testing the subject on sign message meaning for each sign. For example, the answers might be

a. I no longer have to stop.
b. I don't know which approaches have to stop, but I do have to stop.
c. Traffic approaching from the left and the right is not required to stop, but I am.
d. Not certain.

Part II tested the preference among the 11 signs for possible use in advance (i.e., as advance warning) of the intersection. This was followed by the comparative ranking of the top 3 of the 11 signs as first, second, or third choice. Finally, the subjects were given the opportunity to provide comments and suggestions.

Before beginning the laboratory sign evaluation test, each subject was requested to complete a five-question (checkoff) classification questionnaire of age, sex, driving experience, and frequency. The laboratory tests began with draft questionnaires and test questions at the FHWA Turner-Fairbank Highway Research Center. The questions, format, and testing procedure, including developing new and better slides of the candidate signs, were revised and the laboratory tests were administered to the groups indicated in Table 5.

TABLE 5 Summary of Laboratory Test of Candidate Signs

Site	Sign Sequence No.[a]	Question Sets[b]	No. of Subjects
FHWA	1	6	38
U.S. Army reserves	2	2	25
	3	2	20
	4	2	3
Maryland State Police Academy	5	2	37
University of Maryland senior class	6	2	30
University of Maryland health class	7	4	75
Total		20	228

[a]The 11 candidate signs were presented in seven randomized orders.
[b]A total of 53 multiple-choice questions (A through D answers, with one best and at least one correct answer) were developed. These were randomly assembled into 20 different sets of 11 questions each. The numbers in the column give the number of different sets used.

The laboratory testing was accomplished for each group by first briefly describing the general problem of excess stop sign control at intersections. Then the test began by showing a slide of a typical four-way stop-controlled intersection that indicated that the control has been changed, and then showing a slide of the first sign in the 11-sign sequence (at the same intersection) and asking the participants to answer a, b, c, or d. This continued until slides of all 11 signs had been shown.

Results of the sign meaning part of the test were almost identical with the preference (Table 6). Ranking of the signs by percentage of correct answers also revealed signs 2 and 6 (CROSS TRAFFIC DOES NOT STOP) to be the first and second choices, respectively, and sign 5 (NO LONGER 4-WAY STOP) to be the third choice (numbers are as indicated in Figure 5).

The classification questionnaire data were analyzed in two ways. First, the numbers of subjects by age and sex were reviewed to ascertain that there was a representative sample. A total of 102 female and 123 male subjects completed the questionnaire. The number of subjects older than 40 years of age was 30 male and 5 female. Thus the older female driver is not well represented. However, all other age groups are well represented, and this small sample of females older than 40 years is not believed to be a significant bias. The data in Tables 7 and 8 give the results of driving experience and frequency versus age and incorrect answers for female and male subjects, respectively. The average incorrect answer was 21.7 percent for female as compared with 19.1 percent incorrect for male subjects. Of the largest category of female subjects--age 20 to 24, with 5 to 9 years of driving experience, who drive every day--43 subjects had 18.6 percent incorrect answers. This is almost identical to male subjects in the same age and driving experience and frequency category, who had 19.2 percent incorrect answers. It was believed that the sample, when broken down into age, driving experience, and frequency, was too small for any statistical analysis to be undertaken.

TABLE 6 Sign Preference from the Laboratory Comparative Analysis

Sign No.	Choice: First	Second	Third	Weighted Preference[a]: Weight	Rank
1	30	9	15	123	4
2	66	40	15	293	1
3	10	11	17	69	9
4	14	19	29	109	6
5	28	27	16	154	3
6	34	41	28	222	2
7	13	6	12	57	10
8	11	18	17	86	7
9	14	22	26	112	5
10	4	12	9	45	11
11	4	19	20	70	8
Total	228	224	204		

[a]1st choice 3, 2nd choice 2, 3rd choice 1.

TABLE 7 Driving Experience and Frequency Versus Incorrect Answers on Sign Evaluation–Female

Age	Experience (years)	Frequency of Driving	Sample No.	Incorrect Answers: No.	Percent
16-19	3-4	Everyday	1	4	36.4
		3-4 times/week	2	1	4.5
		Total	3		
20-24	3-4	Everyday	8	19	21.6
		3-4 times/week	1	6	54.5
		1-2 times/week	1	2	18.2
		2-4 times/month	1	5	45.5
	5-9	Everyday	43	88	18.6
		3-4 times/week	9	17	17.2
		1-2 times/week	4	20	45.4
		Total	72		
25-29	5-9	Everyday	6	8	12.1
	10-14	Everyday	3	8	24.2
		3-4 times/week	2	9	40.9
		Total	11		
30-39	10-14	Everyday	3	8	24.2
	15-19	Everyday	8	31	35.2
		Total	11		
40-49	15-19	Everyday	1	2	18.2
	20+	Everyday	2	4	18.2
		3-4 times/week	1	2	18.2
		Total	4		
50-59	20+	Everyday	1	8	72.7
Total			102	243	21.7

TABLE 8 Driving Experience and Frequency Versus Incorrect Answers on Sign Evaluation—Male

Age	Experience (years)	Frequency of Driving	Sample No.	Incorrect Answers No.	Incorrect Answers Percent
16-19	3-4	Everyday	2	2	9.1
20-24	1	3-4 times/week	1	6	54.9
	1-2	Everyday	1	1	9.1
	3-4	Everyday	7	22	28.6
		3-4 times/week	2	9	40.9
	5-9	Everyday	18	38	19.2
		3-4 times/week	3	8	24.2
		1-2 times/week	2	10	45.5
		Total	34		
25-29	5-9	Everyday	4	10	22.7
		3-4 times/week	1	0	0
	10-14	Everyday	13	26	18.2
		3-4 times/week	2	4	18.2
		1-2 times/week	1	4	36.4
		Total	21		
30-39	10-14	Everyday	4	4	9.1
	15-19	Everyday	18	38	19.2
		3-4 times/week	3	6	18.2
	20+	Everyday	11	26	21.5
		Total	36		
40-49	15-19	Everyday	1	5	45.5
	20+	Everyday	19	22	10.5
		Total	20		
50-59	20+	Everyday	7	15	19.5
60-64	20+	Everyday	1	0	0
		3-4 times/week	2	2	9.1
		Total	3		
Total			123	258	19.1

In summary, the laboratory sign evaluation test results were consistent with the field experience and literature review, and were in agreement with philosophies of state, county, and municipal officials. The black CAUTION sign on yellow background separated from the black message on white background (CROSS TRAFFIC DOES NOT STOP) is the top candidate as a supplementary sign for safe removal of multiway stop signs. The same top portion CAUTION with the bottom message NO LONGER 4-WAY STOP is a close second preference.

RECOMMENDED PROCEDURE FOR REMOVAL OF MULTIWAY STOP SIGNS

The procedures recommended here were developed based largely on the experiences of traffic and law officials from more than 30 political jurisdictions, the laboratory experiments of supplementary signs, and the results of the field testing of these procedures.

The procedures recommended herein may be applied with slight modification to a situation where the right-of-way at an intersection is reassigned (i.e., stop sign reversals). This action might require the creation of a multiway stop condition and then the removal of the unwarranted stop sign(s).

Each local jurisdiction must determine to what degree the recommended procedures apply for a given intersection. Factors such as community concerns, intersection geometrics, speeds, volumes (vehicular and pedestrian), accident history, and sight distance must be considered.

A decision must likewise be made as to how many intersections are to be converted and when. This is where local politics and economics come into play. If mass removals of stop signs are likely to cause an outpouring of public opposition throughout the entire city or town, then perhaps a neighborhood-by-neighborhood or intersection-by-intersection strategy might be developed. In the second instance, if the nature of the intersections is such that supplementary signs are desirable, the timing of intersection conversions would depend on availability of funds to support materials and labor needed to accomplish the conversions.

There are three phases to the removal of multiway stop signs: the preconversion phase, the actual conversion phase, and the postconversion phase. Following all steps in the procedure will ensure that the conversion will minimize hazard to the driving public.

Preconversion Phase

Conduct Traffic Engineering Studies

Traffic studies should be conducted to determine whether a multiway stop intersection is justified (i.e., that all stop signs at that particular intersection are warranted). The warrants presented in the MUTCD (1) should be used as a basis for determining whether the multiway stop control is justified. If necessary, volume counts should be taken to determine whether the MUTCD warrants are satisfied. In addition, accident records should be checked to determine whether the multiway stop was originally warranted because of accident history at the intersection.

From the very beginning the importance of using appropriate supplementary plates (R1-3) in conjunction with stop signs at multiway stop intersections must be emphasized. The proper use of these supplementary plates (3-WAY, 4-WAY, ALL WAY, and so forth) fixes in the motorist's mind that the intersection is in fact a multiway stop intersection. The absence of these plates at a multiway stop intersection could cause confusion or uncertainty on the part of the motorists, thus resulting in an unsafe and inefficient intersection.

If supplementary plates are not in use at an intersection targeted for conversion, they should be added at least 30 days before the actual conversion. Thereafter, the removal of these plates on the day of conversion will further signal to the motorists that a change has occurred at that particular intersection.

Secure Approval for Stop Sign Removals

Permission should be sought to remove those stop signs determined to be unwarranted. In some instances the local council may have previously delegated the authority for traffic control device installation and removal to the individual responsible for traffic operations. These include the traffic engineer, police chief, director of public works, and so forth. If this is the case, removals are expedited.

Phasing of stop sign removals accomplishes several objectives:

1. Lessons learned at one location can be applied to succeeding locations.
2. Individual neighborhoods can be addressed regarding the stop sign removals as opposed to the entire municipality at once.
3. Supplementary signs can be used during future conversions, thereby reducing the inventory required.

4. Work can be accomplished by existing crews without excessive amounts of overtime.
5. Neighborhoods scheduled for future conversions can witness successful actions elsewhere in town, which will alleviate some of their fears.
6. Local approving officials may find this method more acceptable.

Publicize Planned Multiway Stop Intersection Conversions

The activities associated with obtaining legislative approval (as noted in a previous section) often will serve to publicize planned conversions. Notices to neighborhood residents might be dispatched by using any one or more of several media: newspapers, radio, television, utility bills, flyers, individual letters, and community newsletters.

In addition, notice signs should be posted at the affected intersection to alert the motorists who use the intersection of the impending change. The notice signs shown in Figures 6 and 7 were developed as a result of this study.

FIGURE 6 Notice sign for major approach.

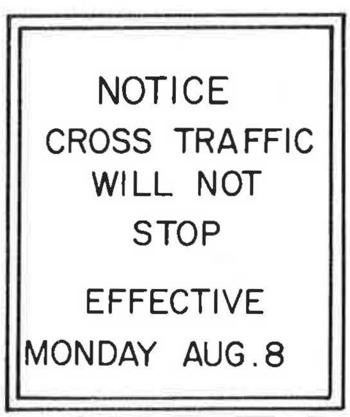


FIGURE 7 Notice sign for minor approach.

Install STOP AHEAD Messages

If not already in use, STOP AHEAD signs (W3-1 or W3-1A) should be installed in accordance with the MUTCD on those approaches that will remain under stop control. These should be considered for use to further emphasize the continued need to stop at a given intersection.

Install Necessary Pavement Markings

If not already present, stop lines and STOP pavement markings may be used in accordance with the MUTCD to highlight the requirement to stop at the intersection.

Conversion Phase

Remove Obsolete Pavement Markings

Before the day of conversion, any stop lines or other pavement markings rendered obsolete by the change should be removed or otherwise obliterated. If rental equipment is involved in the pavement marking removals, several sites should be considered for conversion during the same time period. This would make for more economical and efficient use of rental equipment.

Improve Sight Distance

Sight distance at the intersection should be improved, if necessary, by (a) imposing parking restrictions, (b) pruning vegetation, or (c) adjusting location of stop line. If the stop line is placed directly opposite the stop sign, and the stop sign is placed some distance from the intersecting street's curb line because of intersection geometrics, a motorist could have his sight distance severely reduced. In this instance the stop line should be placed forward of the stop sign but no closer than 4 ft to the intersecting street's curb line. This allows a driver to move to a point of improved visibility from which he can better make gap-acceptance decisions.

Change Signs

The following sequence of events should occur before the beginning of the morning peak period on the day of stop sign removals:

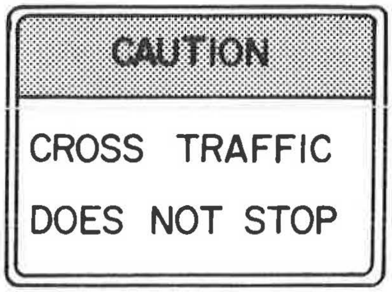
1. On the minor approach replace supplementary plate(s) and sign as shown in Figure 7 with the caution sign shown in Figure 8.
2. On the major approach, after completing the action in 1, remove the unwarranted stop sign(s), supplementary plate(s), and accompanying post(s) and notice sign(s) (Figure 6).
3. Remove unnecessary STOP AHEAD sign(s), including the post(s) on which they are mounted.
4. Replace 24-in. stop signs with 30-in. stop signs for added emphasis. (This could be a temporary or permanent change.)

It is extremely important to convert the intersection before the morning peak period so as not to cause doubt in the motorists' minds concerning the previously publicized action.

Postconversion Phase

Conduct Traffic Engineering Studies

As warranted by the nature of the intersection, any number of studies might be conducted to determine the effectiveness of the stop sign removal action as



Specifications

1. Overall sign dimensions = 24 x 18 in.
2. Caution band dimensions = 24 x 6 in.
3. Lettering height = 4 in.
4. Colors: black letters on yellow and white backgrounds
5. Surface: reflective sheeting

FIGURE 8 CAUTION sign for approach still required to stop after multiway stop intersection conversion.

well as to prepare the traffic engineer to address issues and concerns raised by interested citizens. Typical studies in varying degrees might include traffic volumes, traffic accidents, conflicts, speed, and observance of traffic control devices. These studies are all discussed at length in the Transportation and Traffic Engineering Handbook (8).

Request Police Enforcement

If it is observed or reported that speeding is a problem, increased police enforcement should be requested at the location in question. The length of this increased enforcement would depend on past experiences with similar problems in the local area.

Remove Caution Signs

Ninety days after the intersection has been converted, the caution signs should be removed from beneath the remaining stop signs. If 30-in. stop signs are to be replaced by 24-in. stop signs, that action should be accomplished at this time.

Continue Traffic Engineering Monitoring

After the intersection has been converted, it should be continuously monitored as a part of the regular traffic engineering program. Accident data should be evaluated for the 12-month period following conversion of the intersection to determine whether the modified traffic control condition is adequate. Comments from interested citizens should continue to be received and evaluated during this period.

CONCLUSIONS AND RECOMMENDATIONS

As a result of this study it has become obvious that no uniform procedures exist with which to convert multiway stop-sign-controlled intersections to lesser forms of control with minimum hazard. It is likewise concluded that little documentation is available concerning the actual conversion processes in the various jurisdictions nationally. Some jurisdictions do have complete studies available documenting their actions. The procedures developed in this paper and pilot tested in the field have been shown, through limited testing, to have great potential for minimizing the hazards associated with multiway stop sign removals.

It is recommended that the procedures developed herein be implemented and that the National Committee on Uniform Traffic Control Devices consider the two notice signs and the warning sign (CAUTION, CROSS TRAFFIC DOES NOT STOP) for inclusion in the MUTCD.

Discussion

Bhagwant N. Persaud*

Table 4 of the paper indicates that the total number of accidents at the converted intersections changed from 88 in the year before conversion to 144 in the year after. Ligon, Carter, and McGee, quite rightly, do not base any strong conclusions on these numbers. There is a danger, however, that these numbers could be interpreted as implying that the conversions resulted in a 64 percent increase in accidents. Such an interpretation might conceivably serve as a deterrent to the removal of unwarranted multiway stop control. The danger arises from the fact that, if the conversions were mainly done at intersections that recorded few or no accidents in the before period, as one might suspect, then the observed increase in accidents could be an illusion and an artifact of chance and may not be indicative of a real degradation in safety.

This phenomenon is illustrated by the data in Table 9, which gives changes in numbers of accidents at intersections in Philadelphia that retained multiway stop control during 1973 and 1974. For

TABLE 9 Changes in Number of Accidents at Multiway Stops in Philadelphia

No. of Sites	No. of Accidents per Site 1973	1974	Change (%)
81	0	1.23	Increase
85	1	1.40	+40
58	2	1.60	-20
30	3	1.47	-51
8	4	2.25	-44
7	5	1.71	-66
4	6	1.00	-83
1	7	6.00	-14
1	8	3.00	-63

example, the data in the table indicate that the 85 such intersections that recorded 1 accident in 1973 recorded, on average, 1.4 accidents in 1974, an increase in accidents of 40 percent. As one might expect, those intersections that recorded no accidents in 1973 also experienced an increase in acci-

*Transport Safety Studies Group, Department of Civil Engineering, University of Toronto, Toronto, Ontario M5S 1A4, Canada.

dents. Because the intersections remained unaltered, these increases are a result of chance, or, to use the technical term, regression-to-the-mean. [For further discussion of this phenomenon see the work by Hauer and Persaud (9).]

The 172 intersections examined by Ligon et al. recorded, on average, 0.51 accident in the before period and 0.84 accident in the after period. If some Philadelphia intersections that averaged 0.51 accident in 1973 were converted, then by interpolation from Table 9 these intersections would have recorded, on average, 1.32 accidents in 1974 if the conversions left safety unaffected. In other words, the average number of accidents recorded after conversion would have had to be higher than 1.32 to support a conclusion that removal of multiway stop control leads to a degradation in safety.

Although one might reasonably question whether the Philadelphia intersections are representative of the intersections studied by Ligon et al., this should not detract from the main point of the discussion--that it is misleading to draw conclusions about the safety effect of traffic control measures by simply comparing before-and-after accident records. By deemphasizing the apparent increase in accidents observed in this study, the authors have avoided this pitfall. It is hoped that, after this discussion, others will be persuaded to do likewise.

Authors' Closure

Persaud presents an interesting discussion concerning the safety aspects of the removal of unwarranted stop signs. It was thought that 3 years of accident data before and 3 years after conversion, as well as using control (nonconverted) sites (3 years before and 3 years after) would result in a meaningful experiment, statistically, because of the small number of accidents. Unfortunately, the agencies cooperating in the FHWA study could not provide the 3-year before data base and the study had to be completed in about 1 year. Persaud's Philadelphia example as well as his discussion of accounting for accident change due to chance agree with the authors' intuition and strengthens the recommendation for removal of unwarranted stop signs.

REFERENCES

1. Manual on Uniform Traffic Control Devices for Streets and Highways. FHWA, U.S. Department of Transportation, 1978.
2. W.E. Baumgaertner. In Search of Effective Speed Control. Technical Notes, Dec. 1980, pp. 12-16.
3. R.F. Beaubien. Stop Signs for Speed Control? Traffic Engineering, Nov. 1976, pp. 26-28.
4. G. Funkhouser. Stop Signs Do Not Control Speed. Technical Notes, July 1978, pp. 6-7.
5. K. Parezo. A Study of Stop Sign Control in Residential Areas of Howard County. M.S. thesis. University of Maryland, College Park, 1979.
6. An Evaluation of Unwarranted Stop Signs. San Francisco Bay Area Section, ITE, Feb. 1979.
7. J.E. Upchurch. Development of an Improved Warrant for Use of Stop and Yield Controls at Four-Legged Intersections. Ph.D. dissertation. University of Maryland, College Park, 1982.
8. Institute of Transportation Engineers. Transportation and Traffic Engineering Handbook. Prentice-Hall, Englewood Cliffs, N.J., 1982.
9. E. Hauer and B. Persaud. 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 Traffic Control Devices.

Evaluation of Curve Delineation Signs

BARTON E. JENNINGS and MICHAEL J. DEMETSKY

ABSTRACT

The three post-mounted delineator systems currently used in Virginia were tested at five sites for their effectiveness in controlling run-off-the-road accidents. The changes in speed and lateral placement noted with the systems in place were taken as driver responses to the systems. The study indicated that drivers react most favorably to chevron signs on sharp curves greater than or equal to 7 degrees and to standard delineators on curves less than 7 degrees. It is suggested that statewide use of delineators based on these findings will improve the safety and uniformity in delineation on the rural highway system.

Travel on rural roadways is noticeably different from travel on urban streets. On the former, vehicular speeds are generally higher, the road surface usually is narrower and not as well marked, and the severity of accidents is greater than for urban highways (1).

Several studies have pointed out that a high proportion of the accidents that occur on rural curves happen at night and usually involve a single vehicle that runs off the road (1,2). For a majority of the rural roadways, those with average daily traffic (ADT) of less than 2,100 vehicles, single-vehicle run-off-the-road (ROTR) accidents have been reported to account for more than 40 percent of all accidents, with nearly one-half of these involving a personal injury or fatality (1,2).

Post-mounted delineators (PMDs) of various shapes, colors, and types have been used throughout the United States in an attempt to reduce the number of ROTR accidents. These markers have proved to be effective, especially at night or during adverse weather conditions when roadway markings may be covered (3).

The PMD has been demonstrated to be capable of influencing a driver's judgment of the sharpness of a road curve. This influence can be used to modify the pattern a driver follows through a curve, and thus to promote safety on rural highways (1).

CURRENT PRACTICE

The three basic types of delineation or alignment signs used on rural roadways in Virginia are

1. The 3 x 8-in. reflector on a wooden post (ED-1),
2. The 6 x 48-in. special striped delineator, and
3. The chevron alignment sign (WI-8).

Figure 1 shows these sign types (4).

Two general approaches are used in selecting delineators for a site. The Manual on Uniform Traffic Control Devices (MUTCD) (5) is an often-quoted source for delineation selection for freeways and major roadways. This manual recommends spacing, location, and height for the delineators without recommending the type of delineator to be used. The MUTCD states that "delineation is intended to be a guide to the vehicle operator as to the alignment of the highway; whatever is needed to provide that guidance in a clear and simple way should be installed" (5).

The second method of selection is local practice. A survey of each of the nine operating districts of the Virginia Department of Highways and Public Transportation found wide variations in the use of PMDs, and a review of delineation practices in other states revealed many of the same problems and practices that exist in Virginia. Several states are involved in studies to determine the safest delineation systems for rural roadways. Although the results from the states have not yet been finalized, the following conclusions can be drawn from their data:

1. Large chevrons are not effective and have little effect on speed, braking, or lateral placement within the curve (1).
2. Standard delineators in an MUTCD configuration positively affect speed, braking, and lateral placement and are particularly effective on sharp rural curves (1).
3. Rural curves with PMDs have a much lower nighttime ROTR accident rate than curves of similar characteristics without vertical delineation. Tests have shown the reduction rate to be 50 percent or more (6).
4. Long-term effects of PMDs are much less than the initial effect during the first few weeks. This suggests adaptation by local drivers. Because accidents on rural roads often involve drivers unfamiliar with the roadway geometry, this result does not negate the safety benefits of vertical delineators (1).

During the late 1970s and early 1980s FHWA initiated projects with eight state highway agencies to evaluate the effectiveness of different types of PMDs. The study noted that "it is not possible to state that the installation of post delineators under all conditions will result in a reduction in the number of run-off-the-road type accidents. The data that were collected indicate a trend toward reducing run-off-the-road accidents with the installation of post delineators" (6).

OBJECTIVE AND SCOPE

The purpose of this study was to determine in what areas current practice in the placement of the available types of delineator signs could be improved by providing uniformity. The only focus was the effects of different PMDs on driver behavior. Standard 4-in. pavement markings were in place at all test sites. Selected delineation strategies were evaluated and recommendations were developed for selecting the type of sign best suited for given roadway and environmental conditions, after the decision has been made to use vertical delineation at a site.

RESEARCH METHOD

Performance Measures

Studies on driver reactions to delineation systems placed on roadways generally rely on changes in vehicle movement as indicators of the reactions. The two most obvious changes in movement are vehicle speed and placement. The path a driver takes through a curve is dependent on his perception of the curve and of how best to traverse it. Because this positioning changes as the vehicle moves through the curve, it is desirable to record the placement and speed of the vehicle at several locations during the maneuver (1,7).

Vehicle speed is an indication of the apparent severity of the curvature of the roadway. Slow speeds entering the curve indicate that the driver is aware that the curve exists. Fast speeds at the start of the curve with slower speeds near the middle indicate braking by the driver, probably because the curve is sharper than he perceived it to be. Acceleration in the curve would indicate that the driver perceived the curve to be sharper than it actually is.

The path of the vehicle through the curve is also a good indication of the perceived sharpness. Movement across the centerline may indicate that the curve is not as sharp as it looks. This centerline encroachment may also be caused by objects along the shoulder of the road that the driver perceives to be a threat.

Vehicles traveling close to the right-hand edge of the road may indicate that the curve is sharper than it appears. This occurrence may also be an indication of high ADT, which causes drivers to feel unsafe driving near the centerline (1,7).

Although there are numerous exceptions to these hypotheses, in general it can be stated that a satisfactory delineation system is one that will pro-

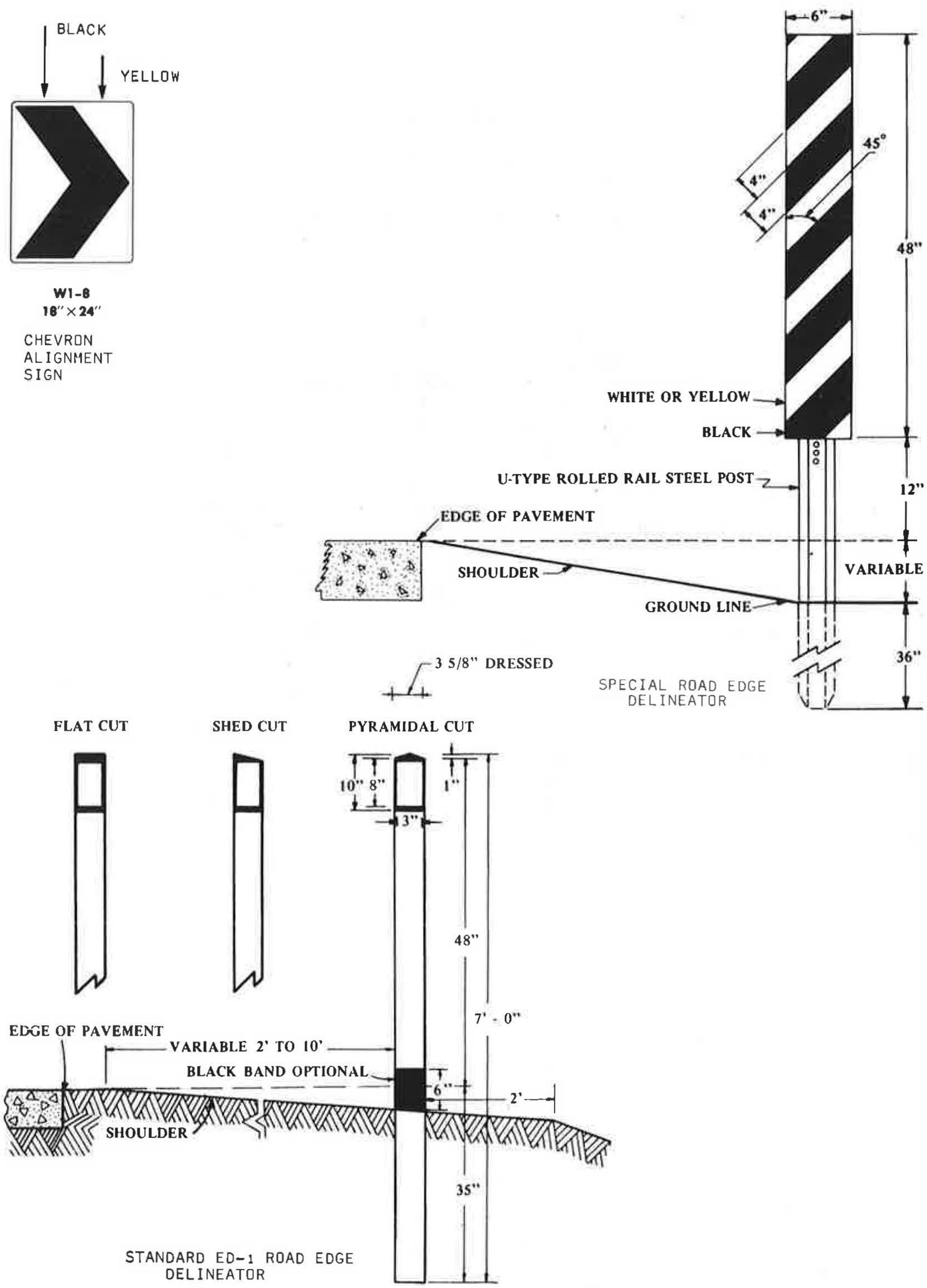


FIGURE 1 Alignment signs tested *(4)*.

duce uniform speeds and placement of a vehicle as it moves through the curve. The system will negate the need for excessive braking in the curve, and the absence of a change in speed when a vehicle is within the curve is a prime indication that the driver of the vehicle has correctly perceived the curvature of the road. Also, it will minimize encroachments on the centerline and edge line, thereby leaving most of the vehicles driving in the center of the lane (1).

On some roads vehicle type could be an important third item that should be recorded. For example, sites should be noted where exceptionally large numbers of heavy trucks are present or where continuous grades reduce the speeds of these trucks but not those of other vehicles. Because large trucks constitute a small percentage of the normal traffic on most rural roads, data for trucks were not studied separately.

Statistical Method

The effectiveness of different delineation treatments was measured by using the chi-square goodness-of-fit test. Here performance data for the marked roadway were compared with those obtained while the curve was unmarked.

The purpose of this analysis was to determine the value of statistical similarity for the delineation treatments of the marked roadway compared with those of the roadway without markers. The larger the value of α that was obtained, the more similar were the data for the two tests. A small value of α indicated that the delineation treatment had significantly altered the driver's path or speed or both through the curve. For example, an α value of 0.10 means a 10 percent level of significance, which in turn indicates a significant change in driver performance in the curve.

The results of this statistical evaluation indicated that there was no significant change in speeds after the delineators were installed. Most values were in the 0.90 range. However, there were significant changes in the lateral placement of vehicles. For this reason, the lateral placement changes were taken as the critical elements in the study; the changes in speed were noted for additional information.

Delineation System and Technique Selection

Delineation systems vary from exotics such as ascending and descending patterns, in-and-out patterns, and sign mix patterns to the more traditional systems currently used in Virginia (1). Because this investigation was intended mainly to test the systems used in Virginia, only three conventional systems were investigated (see Figure 1). The only variation made was that the wooden posts used with the standard road edge delineators were not painted. The decision to use treated but unpainted posts was supported by a study that involved the possible use of untreated posts, which found little difference between visibilities for the two types of posts (1). The MUTCD-recommended spacing and placement for standard delineators was used, as is often done in Virginia.

The most effective placement pattern for chevrons has not yet been determined. Most districts in Virginia use their own judgment to determine the placement and spacing of the chevron signs. The placement varies from a pattern in which one sign is always visible to a pattern in which at least three signs are visible. It was decided that because most of the districts recommend that three chevron signs be in sight such a pattern would be used. In examining MUTCD placement patterns, it was noted that the recommended spacing for standard delineators generally provided that four to six delineators would be in the drivers' view (2). By using this information, it was decided to space chevrons at a distance twice that recommended by the MUTCD for traditional delineators. This spacing proved adequate for this study.

Field Data Collection

To record the speed and lateral placement of the vehicles moving through the curve, a Leupold and Stevens traffic data recorder (TDR) was used. Eight tape switches were used to record data at the beginning and near the midpoint of the curve. The switches were temporarily placed from the edge of the centerline to the shoulders of the road. The leads from the switches were connected to the TDR, which was concealed off the roadway.

The switches were placed on the roadway in a predetermined pattern (Figure 2). The use of 6-ft spacing between matching channels (switches) allowed a variation in placement of 0.75 in. with less than a 1 percent change in speed or lateral placement--an important factor in field installations. As an automobile's tires crossed the first and second switches, their circuits were opened. The third switch closed the first circuit to generate the time from switch 1 to switch 3 and the vehicle's velocity. The fourth switch, which was laid at a 45-degree angle to facilitate field measuring and placement, closed the second circuit to generate the time from switch 2 to switch 4. The placement of the vehicle was then calculated by using the following formula:

$$\text{Lateral placement} = 6 * \text{Tan}\ (\theta)\ [(S_1/S_2) - 1] \qquad (1)$$

where

- 6 = distance (ft) separating the speed detector switches,
- θ = angle of the lateral placement switch = 45 degrees,
- S_1 = speed of the vehicle measured by the speed switch, and
- S_2 = speed of the vehicle measured by the lateral placement switch.

Input from the tape switches was recorded on cassette tapes, and the data were processed on a computer. The output included volume, velocity, and

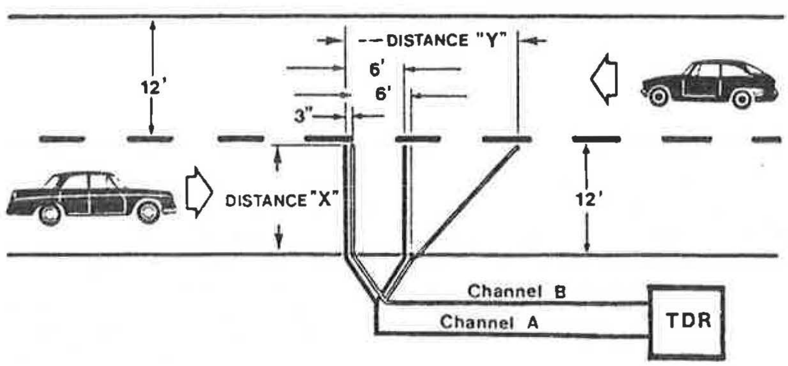


FIGURE 2 Configuration for data collection using two TDR channels per lane.

vehicle type information for the 10 zones into which each lane was divided for lateral placement measures.

Zones 1 through 9 were of equal width, whereas zone 10 represented vehicles that encroached more than 1 ft across the centerline. At the sites tested, zones 1 through 9 were each 8 in. wide (Figure 3).

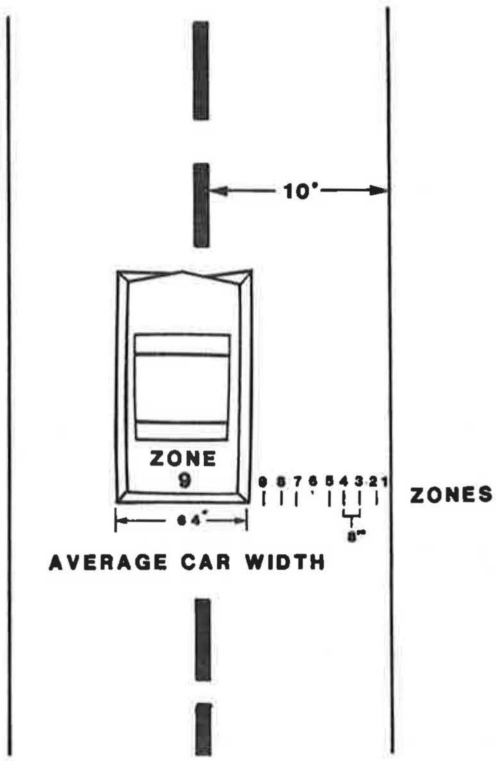


FIGURE 3 Illustration of lane zones.

By using this zonal width, it could be concluded that vehicles in zone 10 represented possible head-on collisions, whereas zones 8 and 9 represented possible sideswipe accidents. Zones 8 through 10 (zones 7 through 10 at the narrowest sites) were considered to be the centerline encroachment zones. Any vehicles in these zones were considered to be candidates for multivehicle collisions.

The data by lane-zone allowed trace data to be determined for average vehicles. This vehicle trace, combined with the velocity averages, was used to determine the effectiveness of the delineation treatments. That the use of average trace data tends to overshadow individual vehicle performance, especially at the two extremes, is of some concern for high velocity areas but is of no concern for low velocity areas (1).

Site Selection

Two groups of roadway sites were used. Sites in the first group were already marked with PMD devices and were used to study the data-collection system as well as to obtain base data (pretest program); those in the second group were initially free of any vertical delineation and were used in the actual testing program. Data were collected once at each pretest site and seven times at each test site. The first collection was taken while the test site was still without markers. Then the site was studied with each vertical delineation device in place to determine short-term effects and it was studied again several weeks later to determine long-term effects.

The following criteria were used to select the sites:

1. Proper signing using current spacing and erection techniques (pretest);
2. No delineation devices (test);
3. No obstacles (driveways, and so forth) on shoulder,
4. Accident history,
5. ADT of 1,000 to 3,000,
6. Location within 1-hr drive of Charlottesville,
7. Rural location,
8. All curves in same construction district to expedite the project,
9. Roadways carry at least some out-of-state traffic, and
10. Standard pavement markings at centerline and edge of pavements.

By using these criteria, a listing of candidate roads and locations was accumulated through interviews with highway officials. Each road or site was then evaluated to determine its suitability for testing.

Technical data for each curve were obtained from the headquarters of the district in which it was located; these were used to group the sites by length of curve, degree of curvature, and degree of grade. The pretest program indicated that vehicle placement was not significantly different in curves with different grades, so grade was not initially considered as a major influence on vehicle placement.

In the field evaluation of a test site, a vehicle was driven through the curve several times, the site was examined for signs of heavy braking or ROTR incidents, and a series of photographs was made. The data in Table 1 identify the sites chosen for the pretest and test phases of the study.

SITE EVALUATIONS

Preliminary Observations

The sites designated 1 through 8 in Table 1 were used for the pretest phase of this study, which was conducted to test the TDR equipment and to determine if the data obtained would allow meeting the study objectives. The data revealed some similarities in driver response characteristics for the different delineation treatments.

As an example, the data in Table 2 give the percentages of vehicle travel and average speeds in each zone across the lane for special pretests at sites 8 and 2. The data are statistically similar for both placement and speed; $\alpha = 0.250$ and 0.950, respectively, which indicates that two sites with different physical characteristics may induce similar driver responses for the same type of delineation signing.

There were also similarities between two of the chevron-marked sites. Here the zonal vehicle placements were not as significantly alike ($\alpha = 0.025$), but the average speed, placement, and centerline encroachment of sites 5 and 7 resembled each other.

Even though these data do not conclusively demonstrate that vehicle paths at sites with the same delineation systems are similar, they do indicate that the patterns are similar at some sites.

In studying the data and site characteristics, it is not the similarity that is worth noting but rather the general trends revealed in the vehicle

TABLE 1 Description of Study Sites

Site No.	Route	County	Curve location	Horizontal Curvature	Radius of Curvature (ft)	Length of Curve (ft)	Lane Width	Grade (%)	1982 ADT	Treatment
Pretest Sites										
1	20	Albemarle	0.6 mile south of I-64	10° 16'	558	237	12 ft 11 in.		3,440	Without markers
2	231	Albemarle	5.0 miles north of VA-22	8°	716	311	9 ft 2 in.	-4	2,600	Existing; special
3	20	Albemarle	7.6 miles south of I-64	8°	716	225	11 ft 0 in.		3,400	Existing; special
4	20	Orange	South of VA-616	4° 30'	1,273	240	9 ft 1 in.		3,990	Existing; special
5	20	Albemarle	Albemarle/Orange county line	12°	447	200	9 ft 0 in.	-5	2,180	Existing; chevron
6	33	Greene	9.6 miles west of US-29	11°	521	323	9 ft 0 in.	-6	2,860	Existing; chevron
7	33	Greene	9.7 miles west of US-29	7°	819	387	9 ft 0 in.	+5	2,860	Existing; chevron
8	231	Albemarle	6.9 miles north of VA-22	4°	1,433	470	8 ft 10 in.	-2	2,600	Existing; special
Test Sites										
9	20	Albemarle	6.8 miles south of I-64	12°	478	215	10 ft 4 in.	+2	3,440	All
10	33	Orange	At VA-652 and VA-664	5°	1,146	824	9 ft 4 in.		3,005	All
11	231	Albemarle	5.5 miles north of VA-22	5°	1,146	748	9 ft 8 in.	+4	2,600	All
12	22	Albemarle	East of VA-783	8°[a]	700[a]	300	9 ft 7 in.	-2	1,530	All
13	208	Louisa	South of VA-642	4°	1,433	583	10 ft 3 in.	-4	2,740	All

[a]Estimated.

TABLE 2 Example of Data Similarities for Sites 8 and 2–Day

	Distribution by Zone in Lane (%)		Avg Zonal Speed (mph)	
Zone	Site 8	Site 2	Site 8	Site 2
Beginning of Curve				
1	1.6	2.4	51.7	49.2
2	13.8	14.2	53.3	52.3
3	26.9	26.4	52.4	51.6
4	29.4	31.9	54.2	53.5
5	17.7	18.4	53.4	52.6
6	5.8	3.9	52.5	54.4
7	3.6	2.1	54.4	54.8
8	0.7	0.3	54.4[a]	54.3
9	0.1	0.3	65.0[a]	57.0
10	0.5	0.3	44.6[a]	53.0
Middle of Curve				
1	0.3	0.6	47.0	48.1
2	3.0	2.7	48.6	48.2
3	6.9	8.5	48.1	49.5
4	22.0	16.0	52.3	52.2
5	24.8	24.5	53.1	52.6
6	19.7	20.4	53.9	53.8
7	16.7	18.8	53.9	54.1
8	4.5	4.5	55.1	55.0
9	1.2	2.6	54.2	57.4
10	0.9	1.5	58.6	57.6

[a]Small number of data points accounts for wide variation in speeds.

data. The consistency in average lateral placement and speed alterations indicates that drivers react in a predictable manner to the different delineation techniques.

The data in Table 3, which give the results of seven tests at the beginning of the curve on site 10 during daylight (6:00 a.m. to 8:00 p.m.), demonstrate how the data can reveal trends in driver reaction. The data are broken down into the 10 zones for each test and include an additional total for possible centerline encroachments. Depending on lane width, the possible encroachments would occur in one of the last three or four zones. At this site, vehicles in zones 7 through 10 experienced encroachments. Centerline encroachments increased during all of the tests.

The percentages given in Table 4 reveal the general trend that vehicles travel away from the edge of the road when delineation signing is in place. The averages and variances in Table 4 more clearly reveal the change. Again, all of the tests of the delineation systems show similar movements, in this case, a strong movement away from the edge line. Also, there was a slight increase in the placement variance, which is used to determine how well defined the new path through the curve is.

The data in Table 5 reveal how vehicle speeds were affected by the new delineator signs. As can be seen, all of the systems induced an increase in speed during the day. The increase in speed with the chevrons was much less; this might indicate that the drivers perceived the chevron signs as obstructions close to the traveled way more than they perceived the other delineators as obstructions. Also, the speed variance increased greatly for the chevron signs whereas speed decreased when other delineators were used. This again points to the possibility that drivers were apparently not as comfortable with the chevrons as they were with the other delineator systems.

Main Tests

Sharp Curves

Two of the five curves studied in depth, sites 9 and 12, are considered to be sharp (curvature greater than 7 degrees). The data from both indicate that the chevron sign is the most favorable form of delineation at these sites. The data for site 9 indicate that of the three delineation systems, the chevrons produced the lowest probable centerline encroachment and, on average, the traveled paths of vehicles on roads with chevrons were closest to being centered in the lane. The placement variability was also lower than that of vehicles driving on roads that have other delineator systems.

The speeds at site 9 also indicated that chevrons performed best on that curve. The average speeds were slightly higher than those of the other systems--a maximum of 2 percent--but the speed variances were among the lowest found.

The data taken at site 12 revealed much the same trends. The centerline encroachment was lower for the roads with chevrons, and the average vehicle path was the most desirable, especially at the middle of the curve where it was about 0.5 ft farther away from the centerline. The placement variance was about average for the three systems studied.

TABLE 3 Example of Vehicle Placement by Percentages, Beginning of Curve (site 10, day)

	Curve Treatment						
		Standard		Special		Chevron	
	Without Markers	Short Term	Long Term	Short Term	Long Term	Short Term	Long Term
Zone							
1	0.1	0.1	0.1	0.1	0.1	2.0	0.1
2	3.4	1.4	0.6	0.8	1.3	2.4	2.6
3	13.7	9.5	5.7	7.1	6.1	7.1	7.9
4	31.6	23.0	22.4	24.5	24.2	21.6	21.8
5	28.5	30.3	29.0	29.1	31.8	28.3	30.7
6	14.9	19.6	24.4	22.5	18.8	20.5	19.9
7	5.8	11.3	13.3	12.0	12.5	13.0	13.2
8	0.9	3.2	3.1	2.5	3.4	3.2	2.7
9	0.8	1.2	0.7	0.6	1.0	1.3	0.8
10	0.3	0.5	0.6	0.9	0.8	0.7	0.4
Possible centerline encroachment	7.8	16.2	17.7	16.0	17.7	18.2	17.1
Total volume for test period	924	862	975	932	912	709	978
Chi-square	–	0.05	0	0	0	0	0.005

TABLE 4 Example of Lateral Placement and Variability Data—Site 10 (ft)

	Beginning of Curve			
	Day		Night	
Curve Treatment	Lateral Placement	Variability	Lateral Placement	Variability
Without markers	2.75	0.75	3.21	0.87
Standard				
Short term	3.08	0.86	3.50	0.86
Long term	3.19	0.80	3.69	1.10
Special				
Short term	3.12	0.82	3.78	1.24
Long term	3.14	0.87	3.79	1.12
Chevron				
Short term	3.08	1.15	3.75	1.23
Long term	3.08	0.86	3.72	1.13

TABLE 5 Example of Vehicle Speed and Variability Data—Site 10 (mph)

	Beginning of Curve			
	Day		Night	
Curve Treatment	Speed	Variability	Speed	Variability
Without markers	51.8	46.2	53.3	41.0
Standard				
Short term	53.0	44.9	52.8	43.6
Long term	53.6	43.6	53.4	46.2
Special				
Short term	53.0	44.9	51.6	47.6
Long term	52.9	42.3	52.6	33.6
Chevron				
Short term	52.1	77.4	51.0	57.8
Long term	51.9	56.3	52.4	54.8

The chevrons at site 12 were not as successful in dealing with speed as they were at site 9. The speeds averaged about 50 mph, which were greater than the 35- and 40-mph speeds recommended by two signs in the area. For the chevrons, daytime speeds were slightly lower than for the other systems, whereas nighttime speeds were greater by as much as 2 mph. The speed variances for the chevrons were also slightly greater during the day, but at night they were about the same as for the other two systems.

Gentle Curves

At sites 10, 11, and 13 the standard and special delineators provided the best delineation; usually the standard delineators were preferred.

At site 10 the standard delineators produced the lowest levels of centerline encroachment and an average lateral placement that was slightly better than those of the special delineators or the chevrons. During the day the chevron produced the best lateral placement. However, at night the lateral placement for the standard delineators, which had a much smaller variance, was the best of the three systems tested.

The speed data for site 10 revealed the special delineator treatment to be superior; the vehicle speeds were greater than average and the speed variance was lower than that of the other two systems. The standard delineator proved to be the second most effective system in terms of speeds and speed variances.

The testing at site 11 indicated that the chevron signs produced some of the lower centerline encroachment figures, especially at the middle of the curve. There was little difference between the standard and special delineator treatments.

In average vehicle placements, no one system appeared to have a major advantage over the others. The special delineators caused the average vehicle path to be slightly closer to the center of the lane than did the other systems. The variance in vehicle placement for the special delineators was also the lowest, which indicated that the delineators were more uniformly accepted at this site.

The speeds recorded at this site changed little from one delineation system to another. The chevron sign produced the slowest speeds, but the speeds were variable. The standard and special delineators produced nearly the same speeds and variances; however, the changes over time for the two system were opposite--the speeds increased for the standard delineators and decreased for the special delineators. For both types of delineators, the speed variances decreased; the special delineators produced the largest decrease.

Site 13 was the most difficult of the test sites to analyze because of the loss of the data for the special delineator short-term test and the repaving of the roadway before the chevron long-term test.

Vehicle placement and speed data, however, do reveal that the standard delineators produced the lowest levels of centerline encroachment. Use of standard delineators also resulted in low levels of vehicle placement variance and produced a vehicle path near the center of the traveled lane through the curve. The placement variance results for the special delineators indicated that they are more effective in producing uniform traffic movements at night than the standard delineators, but the average lateral placement of vehicles traveling on roads with special delineators was much closer to the centerline. The chevron signs produced the highest variances in placement, but lateral placement was similar to that of the standard delineators.

The chevrons did a much better job, judging from the average speeds. They induced the lowest variance of speed of all of the systems at site 13. The standard delineators revealed good variances during the day but had the largest ones during the night; they also produced the lowest speeds during the day and the highest at night.

The results for the special delineators were always satisfactory, but never the best. This may indicate that, if only one type of delineation treatment is to be used in the state, special delineators would be the most appropriate because they produce no extreme changes in vehicle paths while still providing suitable guidance through the curve.

This general trend revealed by the data (preference for standard or special delineators for these less sharp curves) follows the guidelines that most of the districts in the state use. Therefore, it would appear that the use of these two signs is correct for those sites with a curvature of less than 7 degrees.

Discussion of Findings

All of the Virginia highway districts follow the MUTCD spacing guide for standard and special delineators, so the only problem found in the state related to spacing was with the chevron signs. This project used the system practiced in West Virginia; that is, the regular MUTCD spacing was doubled for the chevron signs (6). In the tests conducted this spacing proved to be successful in providing guidance without using an excessive number of signs.

By using the data and associated inferences obtained from the field tests along with information obtained in the survey of state delineation practices, a simplified delineation policy can be developed. For moderate curves (less than 7 degrees), where delineation is deemed to be necessary, the use of standard delineators spaced as recommended by the MUTCD appears to be the most satisfactory choice. This choice does present some problems to the state, the most significant of which is that the Salem, Suffolk, and Northern Virginia districts reported no current use of these delineators. Another problem is that many such curves are marked in other ways. However, this should be of little concern because the use of delineators already varies from site to site. The use of only standard delineators will eventually result in a more uniformly delineated highway system.

Previous studies tended to question the acceptability of chevron signs. They generally have reported that the signs induce an excessively large number of centerline encroachments along with little, if any, change in vehicle speeds (6). This was not found to be true at all of the five sites studied in this project. Chevrons produced less centerline encroachment than the standard or special delineators while still providing smaller vehicle placement variances at the sharper curves. Likewise, speeds were also decreased in these curves.

These data are supported by recent studies on the use of chevron signs. A possible explanation for this change in driver reaction is that when the first tests were performed chevron signs were a new delineation technique. Many drivers had never seen the signs before and were confused as to their meaning. With chevrons gaining wide acceptance, drivers are more familiar with the signs and are now capable of interpreting their meaning.

A second factor, and possibly a more important one, is chevron sign spacing. When first used, chevrons were used much as a normal delineator would be. This close spacing and large sign size combined to form a wall-like effect alongside the roadway. Drivers tended to move away from this effect and over the centerline. Spacing the chevrons at twice the normal distance tends to eliminate the wall effect while still providing guidance through the curve.

RECOMMENDED GUIDELINES

Many of Virginia's highway districts have been moving toward the use of different delineator systems for sharp and gentle curves, and this policy is supported by the findings of this study.

It has been determined in this study that drivers do react to vertical delineation along the roadway and that this reaction is related to the layout of the curve. Delineation systems used in curves should be matched to the expected driver responses based on such factors as the curvature of the road and sight distance. To ease this decision-making process, the following recommended guidelines are offered for curves deemed to require delineation because of the degree of curvature and not because of other factors such as the presence of intersections or hazards on the roadway shoulder.

For curves less than or equal to 7 degrees, the use of standard edge delineators (ED-1) is recommended. The spacing should conform to that given in Table 6 (4). The height of the delineator post should be 4 ft above the right edge of the pavement and the post should be located 6 to 8 ft from the edge of the pavement (5).

For curves greater than 7 degrees, the use of chevron alignment signs (WI-8) is recommended. These signs should be erected 6 to 8 ft from the edge of the road at a top-of-the-sign height of 4 ft above

TABLE 6 Suggested Spacing for Highway Delineators on Horizontal Curves *(4)*

Radius of Curve (ft)	Spacing on Curve for Standard Delineators, S (ft)	Spacing on Curve for Chevron Signs, C (ft)
50	20	40
150	30	60
200	35	70
300	50	100
400	55	110
500	65	130
600	70	140
700	75	150
800	80	160
900	85	170
1,000	90	180

Note: The distance is in feet rounded to the nearest 5 ft. Spacing for specific radii not given may be interpolated from the table. The minimum spacing should be 20 ft. The spacing on curves should not exceed 300 ft. In advance of or beyond a curve, and proceeding away from the end of the curve, the spacing of the first delineator is 2S, the second is 3S, and the third 6S but not to exceed 300 ft. S refers to the delineator spacing for specific radii computed from the formula $S = 3 \times (R-50)^{1/2}$. The spacing of chevron signs should be twice that used for standard highway delineators. C refers to the chevron spacing for specific radii computed from the formula $C = 7 (R-50)^{1/2}$.

the right edge of the pavement. The chevrons should be spaced twice the distance of the standard delineators, as noted in Table 6.

The purpose of this study was to compare the existing PMD systems with one another. However, now that it has been revealed that delineation signing can alter a driver's path through a curve, the most effective pattern should be developed. Testing in this area has already been carried out, but the results of these studies have been mixed, with some spacing and height changes showing improvements (7,8).

ACKNOWLEDGMENTS

This study was financed with highway planning and research funds administered by FHWA.

The authors thank Steve R. Blackwell, John D. Shelor, Clyde Giannini, and others from the Virginia Highway and Transportation Research Council who assisted in collecting the field data. Appreciation is also due Jon DuFresne and Fred James of the Culpeper District for their time and assistance in providing information on possible test sites.

Harry Carpenter and personnel of the Traffic Section of the Culpeper District are thanked for their assistance in fabricating and installing the delineation signs. The assistance of Lewis Woodson, Philip Harris, and Mitsuru Saito with the TDR and its programs is also acknowledged and appreciated.

REFERENCES

1. T.H. Rockwell and J.C. Hungerford. Use of Delineation Systems to Modify Driver Performance on Rural Curves. Final Report EES 567. Systems Research Group, Ohio State University, Columbus, Aug. 1979.
2. T.R. Neuman, J.C. Glennon, and J.B. Saag. Accident Analyses for Highway Curves. In Transportation Research Record 923, TRB, National Research Council, Washington, D.C., 1983, pp. 65-69.
3. R.N. Schwab and D.G. Capell. Is Delineation Needed? ITE Journal, May 1980.
4. The Virginia Supplement to the Manual on Uniform Traffic Control Devices for Streets and Highways. Division of Traffic and Safety, Virginia Department of Highways and Transportation, Richmond, Nov. 1980.
5. National Joint Committee on Uniform Traffic Control Devices. Manual on Uniform Traffic Control Devices for Streets and Highways. FHWA, U.S. Department of Transportation, 1978.
6. C.W. Neissner. Post Mounted Delineators. Report FHWA-TS-83-208. FHWA, U.S. Department of Transportation, July 1983.
7. T.H. Rockwell, J. Malecki, and D. Shinar. Improving Driver Performance on Rural Curves Through Perceptual Changes--Phase III. Final Report EES 428. Engineering Experiment Station, Ohio State University, Columbus, March 1975.
8. J.C. Yu and A.C. Arnn. Roadside Delineation Concepts: A National Study. In Highway Research Record 440, HRB, National Research Council, Washington, D.C., 1973, pp. 57-68.

The opinions, findings, and conclusions expressed in this paper are those of the authors and not necessarily those of the sponsoring agencies.

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

Abridgment

Active Advance Warning Signs at High-Speed Signalized Intersections: A Survey of Practice

RONALD W. ECK and ZIAD A. SABRA

ABSTRACT

A synthesis of current traffic engineering practice relative to accident countermeasures at high-speed signalized intersections is presented. The synthesis was prepared by using two main sources of information: a review of published and unpublished literature, and results of a questionnaire survey sent to practicing traffic engineers. Sixty-five state and local agency traffic engineers from all regions of the United States and Canada responded to the questionnaire. Physical environments known to cause problems at high-speed signalized intersections are described. The three most commonly used types of active advance warning devices are discussed along with quantitative and qualitative assessments of their effectiveness. The three types of devices are flashing RED SIGNAL AHEAD signs, PREPARE TO STOP WHEN FLASHING signs, and flashing strobe lights. Active devices are usually installed only as a last resort where conventional countermeasures have not proved to be effective. Although specific situations in which each type of device has been effective or in which its use should be avoided were identified, there is a need for guidelines to define the use, design, installation, and timing of active warning devices.

Introduction of signalization of high-speed highways (defined as approaches with posted speeds greater than or equal to 45 mph) creates the potential for a significant increase in traffic accidents. Two common problems at such locations are the creation of a decision zone and the existence of geometrics such that the signal is not expected or that the display cannot be seen in time.

The most promising countermeasures for this problem appear to be active advance warning devices. These are traffic control devices, placed at or in advance of high-speed signalized intersections, that alter the information provided to drivers on the basis of whether drivers should stop or proceed. With active devices, accident potential is reduced because drivers have additional information on which they can decide the best course of action.

The two principal types of active devices are the . . . WHEN FLASHING signs and signal head supplements. The . . . WHEN FLASHING signs include devices placed in advance of the intersection, which indicate to drivers whether to stop or proceed. Such devices usually take one of two forms, the flashing RED SIGNAL AHEAD sign and the PREPARE TO STOP WHEN FLASHING (or similar message) sign. The flashing RED SIGNAL AHEAD sign displays two messages: RED SIGNAL AHEAD or SIGNAL AHEAD. The neon RED SIGNAL AHEAD message is activated near the end of the green interval or during the yellow interval and remains on throughout the red interval. The word RED flashes alternatively with the words SIGNAL AHEAD. At all other times during the cycle length, the SIGNAL AHEAD message is displayed (nonflashing).

There are many variations of the PREPARE TO STOP WHEN FLASHING signs. Essentially, the device consists of a sign panel with a word or symbol message and yellow flashers that illuminate a predetermined time before the start of red. The signs are characterized by their variety; messages currently in use include STOP AHEAD WHEN LIGHTS FLASHING, STOP ON SIGNAL WHEN LIGHTS FLASH, and BE PREPARED TO STOP.

The strobe light is a flashing white light that supplements the red indication of a traffic control signal. The flashing strobe is intended to draw motorist attention to the signal in situations in which the signal is unexpected or may be difficult to see because of other lights and signs. The flash rate of strobe lights is usually 90 or more flashes per minute; the pulsating strobe appears only with the normal steady red indication.

Although the availability of these solutions has reduced some of the problems, accidents at high-speed signalized approaches appear to be a persistent concern nationwide. To provide guidance to traffic engineers who face decisions about countermeasures at high-speed signalized intersections, there is a need for comprehensive review and evaluation of such countermeasures, both successes and failures.

A research project was undertaken to provide the review and evaluation of such countermeasures. Objectives of the study were

1. To review current traffic engineering practice relative to accident countermeasures at high-speed signalized intersections through (a) review of published literature and (b) a survey of practicing traffic engineers, and
2. To prepare a synthesis of practice on approaches to the problem.

The countermeasures described were included on the basis of having been identified as specific treatments for high-speed intersections. Therefore, many of the more traditional signalized intersection countermeasures such as improved intersection geometrics, left-turn lanes, advance rumble strips, all-red clearance intervals, and flashing signal operation were not considered.

SURVEY OF PRACTICE

To obtain information not available in the published literature, a survey of practicing traffic engineers was conducted to collect data on the nature of the specified accident problem and on the evaluation of appropriate corrective treatments. The questionnaire was sent to engineers with responsibilities in planning, design, and installation of traffic signals. Of the 211 agencies sent questionnaires, 110 responded, although only 65 actually returned completed questionnaires. Response rates were 62 percent for state agencies, 19 percent for local agencies, and 30 percent overall.

ANALYSIS OF QUESTIONNAIRE RESPONSES

Safety and Operational Problems

The first question on the survey form attempted to define circumstances under which safety and operational problems are experienced at high-speed signalized approaches. Respondents were asked to identify circumstances most relevant to their jurisdiction by ranking them (1 = most pressing problem, 2 = next most serious problem, and so forth) and then listing particular safety and operational problems associated with the given circumstance. For state agencies, rural expressways where signalization was unexpected ranked as the most pressing problem. This was followed, in order, by intersections hidden by horizontal curves, rural expressways with heavy truck traffic, and intersections hidden by crest vertical curves. Steep downgrades and intersections hidden by other features were not noted as serious problems by state agencies.

Responses of local agencies were different from those of the states in that only two circumstances stood out as important: intersections hidden by horizontal curvature and those hidden by vertical curvature. In all other circumstances the number of respondents not indicating that a circumstance was a particular problem exceeded those who indicated a ranking for that circumstance.

The relative importance of safety and operational problems associated with intersection circumstances was tabulated separately for state and local agencies. Because there was little difference between the two responses, it was decided to combine the results. Rear-end accidents were the most frequently mentioned accident type overall. This accident type had the highest frequency in all circumstance categories except one (rural expressways with heavy truck traffic). Second and roughly equal in importance with 26 percent of the citations were right-angle accidents and red violations. Within these overall categories, right-angle accidents had a tendency to be associated with limited sight distance, whereas red violations were most frequently mentioned with rural expressways and steep downgrades.

Countermeasures

Several questions on the survey form inquired about countermeasures implemented to reduce problems at high-speed signalized intersections. Results indicated that traditional approaches to intersection accident problems, detector placement, and yellow time adjustment are by far the most frequently used countermeasures at both the state and local levels. At the state level, the flashing RED SIGNAL AHEAD sign was the most widely used dynamic device, with more than 300 installations nationwide. More than 200 PREPARE TO STOP WHEN FLASHING signs were reported, with more than one-half of these being of the ground-mounted type. Only 12 strobe installations were reported nationally. This is significantly less than the number reported in another recent survey (1).

The basis for installation of various countermeasures used by state and local agencies was also reported. Overall, for both state and local agencies, rear-end and right-angle accidents and red violations were the most frequently cited bases for installation, accounting for almost 60 percent of the responses. In general, all countermeasures revealed a similar pattern for basis for installation. The one exception was the flashing SIGNAL AHEAD sign. Both state and local agencies indicated that one of the main reasons for installing the flashing SIGNAL AHEAD sign was as a response to speed problems.

Although results indicated that truck accidents were not a significant problem, a question on the survey form asked whether trucks were given special consideration in countermeasure selection. Of the 22 state agencies responding to this question, 59 percent indicated that trucks were given special consideration. In contrast, only 36 percent of the 18 responding local agencies answered the question affirmatively.

Agencies were asked to provide information on the interval of the signal cycle in which the dynamic devices are activated. In general, the results agreed with what had been expected. For state agencies, roughly two-thirds of the dynamic devices (excluding flashing strobes) were activated at a predetermined time before the start of red (i.e., during green). This percentage was lower for local agencies for which a higher proportion of devices were activated at the beginning of yellow.

Respondents were asked to provide two types of countermeasure cost information: (a) typical installation cost per intersection approach, and (b) annual maintenance cost per approach. For local and state agencies, yellow time adjustment had both the lowest median installation cost and the lowest median annual maintenance cost of all countermeasures considered. The overhead PREPARE TO STOP WHEN FLASHING sign was the most expensive to install, costing about $5,000 per intersection approach. At the state level, ground-mounted PREPARE TO STOP WHEN FLASHING signs and flashing RED SIGNAL AHEAD signs had approximately the same mean installation costs, around $2,500. These responses differed from those of local agencies for which costs for detector placement and flashing RED SIGNAL AHEAD signs were significantly higher than those reported by state agencies. But because of the very small sample size, little confidence is placed in the cost data for flashing strobe lights.

As just mentioned, yellow time adjustment had the lowest median annual maintenance cost of all countermeasures. Note that local agency data will not be discussed here because of limited sample size. The most costly devices to maintain were the overhead and ground-mounted PREPARE TO STOP WHEN FLASHING signs. In the intermediate cost range were detector placement, the flashing RED SIGNAL AHEAD sign, and the flashing SIGNAL AHEAD sign.

Effectiveness Evaluation

Two forms of countermeasure effectiveness evaluation--a subjective assessment and an objective quantitative evaluation--were sought on the survey form. The first question asked respondents to give an overall assessment of countermeasure effectiveness on a scale of 1 (no effect) to 5 (excellent).

In several instances state and local agencies differed in their assessments of countermeasure effectiveness. At the state level, detector placement was rated most effective followed by the flashing RED SIGNAL AHEAD sign. Both countermeasures had ratings of "good" or better. Ranking somewhat lower in effectiveness was a group of three countermeasures (overhead and ground-mounted PREPARE TO STOP WHEN FLASHING signs and yellow time adjustment). Two other devices (flashing SIGNAL AHEAD sign and flashing strobe lights) rated just better than neutral.

In contrast, local agency engineers thought the overhead PREPARE TO STOP WHEN FLASHING sign was the most effective. Three devices were tied as the second most effective countermeasures (ground-mounted PREPARE TO STOP WHEN FLASHING sign, flashing RED SIGNAL AHEAD sign, and flashing strobe lights). The flashing SIGNAL AHEAD sign, detector placement, and yellow time adjustment received the lowest effectiveness ratings. In general, local agencies rated the dynamic devices more highly than did state agencies.

The second question relative to countermeasure effectiveness evaluation asked agencies to include results of any formal studies conducted to evaluate the effectiveness of the four dynamic countermeasures identified on the form. Relatively few agencies had conducted formal studies to evaluate effectiveness. Six state agencies and one local jurisdiction sent copies of reports documenting results of evaluation studies. The amount of data furnished was not sufficient to permit statistical analysis.

It is interesting to note that more than one-half of the studies involved evaluation of the flashing strobe light (1-3). Overall, based on a relatively small number of intersections, there was no clear consensus on the effectiveness of strobes. Strobes appeared to be effective in reducing right-angle and total accidents, but in most cases there were no statistical differences in the number of accidents before and after installations of strobe lights.

An Ohio Department of Transportation before-and-after study (4) evaluated the PREPARE TO STOP WHEN FLASHING sign at six locations. High-speed approaches revealed a statistically significant accident reduction for total, rear-end, property-damage-only, and truck-at-fault accidents.

Maryland evaluated the flashing RED SIGNAL AHEAD sign through a before-and-after study that involved 22 intersection approaches (5). The RED SIGNAL AHEAD sign was determined to be successful in reducing right-angle accidents at sight-obstructed signalized intersections. The device appeared to be more effective in reducing rear-end and total accidents on horizontal curve approaches than on steep vertical approaches.

CONCLUSIONS AND RECOMMENDATIONS

The research described in this paper involved an assessment of active advance warning devices and other accident countermeasures at high-speed signalized intersections. Output from two of the project activities--a literature review and a survey of current practice--was combined to achieve a synthesis of practice (6) on active warning devices.

Both the literature review and the survey of practice indicated that hidden intersections and rural expressways where signals are unexpected are the two circumstances creating problems at high-speed signalized intersections. At such locations, rear-end accidents are the most pressing problem, following by right-angle accidents and red violations. Only when conventional countermeasures such as detectorization or continuously flashing SIGNAL AHEAD signs fail to solve the problem will agencies turn to dynamic devices. When active devices are used, they are installed selectively so that their effectiveness is not diminished by overuse.

The most popular dynamic devices are the flashing RED SIGNAL AHEAD sign, the PREPARE TO STOP WHEN FLASHING sign (and its variations), and flashing strobe lights. Some agencies tend to favor one dynamic device more than others. This may be due, in part, to topography, past experience with the device, and installation and maintenance costs. Of the three devices, flashing strobes have the lowest costs. It was concluded that, in general, the flashing RED SIGNAL AHEAD sign was the most effective dynamic device; traffic engineers gave it a "good" rating. Flashing strobe lights were the least effective of the three active devices; engineers rated strobes closer to neutral than to good in terms of effectiveness. For dynamic devices in general, it was concluded that activation of flashing near the end of green is more effective than activation at the beginning of yellow.

Although this study has identified certain situations in which each type of dynamic device is effective or in which its use should be avoided, there are no general warrants or guidelines for the use of active warning devices at high-speed signalized intersections. An application standard is needed to define the use, design, installation, and timing of active devices. Additional investigations using a combination of field studies and laboratory experimentation with a driving simulator are recommended.

ACKNOWLEDGMENTS

This paper is based on research sponsored by the West Virginia Department of Highways (WVDOH) in cooperation with the FHWA, U.S. Department of Transportation. Appreciation is expressed to the state and local highway agency engineers who took time to respond to the survey of practice questionnaire. Ken Kobetsky and Ray Lewis of WVDOH deserve thanks for supplying data and for providing feedback on project activities.

REFERENCES

1. T.A. Ryan. Strobe-Supplemented Red Signal Indications. In Transportation Research Record 956, TRB, National Research Council, Washington, D.C., 1984, pp. 22-24.
2. W.J. Styles. Evaluation of the Flashing Red Strobe Signal. Bureau of Traffic Projects, Maryland Department of Transportation, Baltimore, June 1982.
3. T.G. Swenson. Use of Strobe Light in Traffic Signals. Final Report. Kansas Department of Transportation, Topeka (undated), 11 pp.
4. R.L. Baker, D. Clouse, and D. Karr. Evaluation of the Prepare to Stop When Flashing Sign. Ohio Department of Transportation, Columbus, Feb. 21, 1980, 7 pp.
5. W.J. Styles. Evaluation of the Flashing Red Signal Ahead Sign. Bureau of Traffic Projects, Maryland Department of Transportation, Baltimore, June 1982.
6. R.W. Eck and Z.A. Sabra. Accident Countermeasures at High-Speed Signalized Intersections: Synthesis of Practice. Department of Civil Engineering, West Virginia University, Morgantown, Aug. 1984, 93 pp.

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

Traffic Circles for Residential Intersection Control: A Comparison with Yield Signs Based on Seattle's Experience

G. SCOTT RUTHERFORD, ROBERTA L. McLAUGHLIN, and EDWIN von BORSTEL

ABSTRACT

The city of Seattle has used more traffic circles than yield signs or stop signs to control four-way residential intersections in the past 10 years. The purpose of these traffic circles is to respond to above-average accident problems at neighborhood intersections without having to rely on the use of yield signs and stop signs. It was found in this study that both traffic circles and yield signs can reduce the number of intersection collisions by about 77 percent. Information reviewed for this study indicates that traffic circles can reduce midblock speeds by significant amounts. Locations with traffic circles reveal a variety of changes in volume after a circle has been placed. These changes are likely to be caused by other conditions in the neighborhood rather than by the circle. A total volume decrease of 2 percent (an insignificant change) was measured for 20 traffic locations. Similar data were not available for stop-sign and yield-sign locations. The cost for a traffic circle is much higher than for a yield sign, but if a city is willing to incur additional costs, circles can reduce the proliferation of traditional control devices, perhaps enhancing the effectiveness of signs elsewhere.

The Seattle Engineering Department is using traffic circles in residential intersections in which accidents occur but usually do not occur enough to warrant the installation of yield signs or stop signs according to criteria used by the city of Seattle.

Examples of traffic circles installed in Seattle's neighborhoods are shown in Figure 1. These devices are simple, round-raised islands placed in the middle of intersections. Circle details are shown in Figure 2. Seattle had approximately 150 intersections with traffic circles at the end of 1983. Although other cities throughout the United States have used an occasional traffic circle in residential intersections, Seattle has developed the most extensive system of traffic circles in the country. In the past 10 years there have been more residential intersections in Seattle equipped with traffic circles than intersections equipped with new yield signs or stop signs.

The rules of the road at intersections with traffic circles are the same as at any other unsignalized intersection. The driver on the right has the right-of-way. Cars turning left may turn left in front of the circle or go around it counter-clockwise. If the situation warrants a change from this then KEEP RIGHT signs are installed.

The purpose of this study was to make quantitative and qualitative comparisons between traffic circles and yield signs. A group of 14 traffic circle locations in Seattle was studied in 1980 (1). The results indicated that the installation of traffic circles had reduced the number of accidents more than 90 percent. However, there was no comparison of a similar reduction caused by the installation of yield signs. At the time of this 1980 study, there were not enough data to determine the effects of traffic circles.

SEATTLE'S POLICIES FOR NEIGHBORHOOD TRAFFIC CONTROLS

To respond to a request or concern about a neighborhood traffic problem, a member of the Seattle Engineering Department staff will look at the accident records for the past 3 to 4 years to determine the number and type of collisions occurring at the location under investigation. If numerous accidents have been reported in the past 3 years, then a visit to the site usually follows.

A site visit might reveal problems such as overgrown vegetation, cars parked too close to the corner, or other problems that may be corrected by

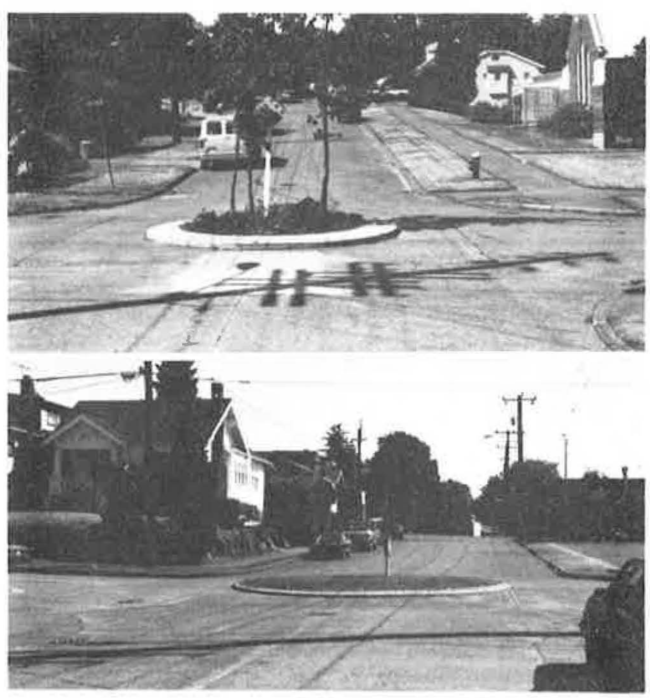


FIGURE 1 Example of traffic circles in Seattle.

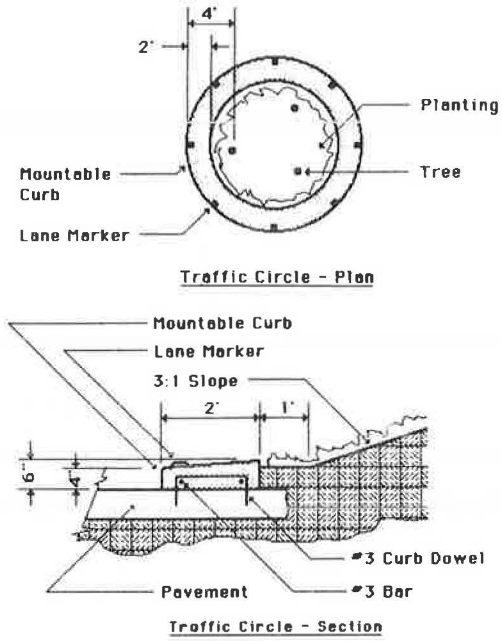


FIGURE 2 Traffic circle details.

some type of action other than installing additional control at the intersection. If there is nothing that is obviously causing problems, then a traffic circle may be recommended.

The procedure used to determine if a circle will be installed requires the neighborhood's support in the form of a petition or mail-back survey to show a majority vote in favor of placing a circle at a particular location. Once this support is shown, additional information is then collected to determine the number of reported accidents over the past 3 years, to measure the 85th percentile speed on one midblock section of the street next to the intersection, and to count the number of vehicles using the major street. This information is used to rank locations in order of problem severity.

The Seattle Engineering Department attempts to put controls that cause the least amount of delay and restriction needed to reduce accidents at four-way residential intersections. Other cities will install yield signs or stop signs at locations based on volumes, sight distance, or number of accidents (2,3). As a result, these cities eventually will have almost every four-way intersection controlled by stop signs or yield signs. It is practices such as these that have caused a widespread use of yield and stop controls at four-way residential intersections.

Rather than reduce the number of accidents to the level required to install yield or stop signs in order to address problem intersections, Seattle has chosen to use traffic circles as a control device that helps prevent accidents from occurring at four-way residential intersections. The use of traffic circles has also provided additional benefits such as significant speed reductions and the ability to respond to the concerns of citizens about traffic safety without having to use yield signs and stop signs.

DATA COLLECTED

Data used in this study were found in the existing files of the Transportation Division of the Engineering Department. The data collected were analyzed for groups of locations that have the same type of control device. Differences between intersections controlled with traffic circles and those controlled by yield signs are determined quantitatively when sufficient before-and-after data exist for each group of locations.

Selection of Intersections

The intersections studied in this project had the following characteristics:

1. Four-way local access street intersection,
2. Primarily residential land use,
3. Change in control between January 1, 1974 and December 31, 1983.

A list of locations that had received new yield signs within the past 10 years was generated from the Engineering Department computer records for traffic signs. Locations that did not meet the criteria previously listed were removed from the list.

Information for traffic circle locations was found in records kept by the Engineering Department. A file is maintained for each intersection. A master list of all circle locations was compiled from these records.

Accidents

A reportable accident in Seattle is defined as a collision that causes $300 or more damage. Information for these accidents is kept on computer tapes and is available for accidents that have happened since January 1, 1974.

Because of the varying installation dates, before-and-after study periods for every location could not be the same. Therefore, to study locations with a reasonable amount of before-and-after accident data, locations studied were chosen for which at least 3 years of before data and 3 years of after data were available.

Speeds

The Engineering Department collects speed information near each intersection before a traffic circle is installed. This information is evaluated to rank intersections that are to receive circles the following year.

After a circle has been in place for at least 6 months, a second speed survey is done at the same location. No studies have been conducted in Seattle to determine the effects of yield signs on speeds.

Volumes

Seattle collects volume information to help assign priorities to locations proposed for traffic circles. An automatic counter is put across one of the legs of the intersection for 7 consecutive days. An average weekday traffic (AWDT) value is determined for the volumes counted on the 5 weekdays.

The volumes are usually measured on the street with the higher amount of traffic. The higher volume street is chosen based on previous short-term counts or from information gathered from nearby residents.

After a circle has been in place for several months, a second automatic count is taken at the same location as the previous count. The before-and-after AWDTs are then adjusted with a seasonal factor to eliminate variations between months of a particular year. The adjusted AWDTs are then compared to determine any changes in volumes occurring on the higher volume street.

DATA ANALYSIS AND RESULTS

Intersection Accidents

To compare the accident experiences at controlled four-way residential intersections, accident totals were generated over the same 10-year period using the following control devices:

1. Yield signs--a group of 65 intersections at which yield signs have been installed between January 1, 1974, and December 31, 1978; and
2. Traffic circles--a group of 38 intersections at which traffic circles were installed between January 1, 1971, and December 31, 1980.

The accident totals were used to determine the average number of reported collisions occurring at these types of intersections during each year from 1974 through 1983. The compilation of accident data over 10 years resulted in the following observations from 1974 to 1983 in Seattle:

1. Uncontrolled four-way intersections averaged about 0.5 accident per year,
2. Yield-sign-controlled intersections averaged between 0.8 accident per intersection each year after 1979 when all 65 locations had yield signs in place, and
3. Traffic-circle-controlled intersections averaged about 0.1 accident per intersection each year after 1980 when all 34 locations had circles in place.

These figures can only be used for trend comparison because the volumes are vastly different for each type of control device.

By reviewing the policies that Seattle uses to place the control devices being discussed, it is expected that proposed locations for yield signs will have higher accident averages before being controlled than intersections selected for traffic circles. To determine those averages and the reductions caused by using these controls, the following data were analyzed: (a) before-accident averages for 1, 2, and 3 years before installing each device, and (b) after-accident averages for 1, 2, and 3 years after each device was installed.

Intersections that had yield signs and traffic circles installed between January 1, 1977, and December 31, 1980, were used for this portion of the study. There were 41 yield-sign intersections and 40 traffic-circle intersections that met these criteria. The results of these before-and-after comparisons are shown in Figures 3 and 4. Results of use of both devices show a 77 percent reduction in number of accidents.

Midblock Accidents

Accidents occurring in all of the four approaches were tabulated for various locations that have the two types of controls. This information was averaged over various before-and-after time periods determined by installation date of the control device at the intersection. The results of this data analysis are given in the following table (note that the data

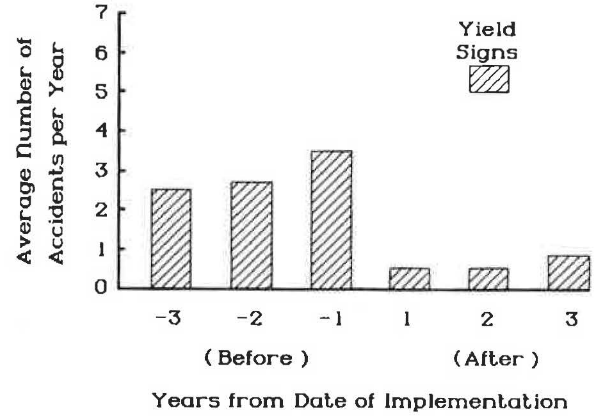


FIGURE 3 Before-and-after accidents at locations using yield signs.

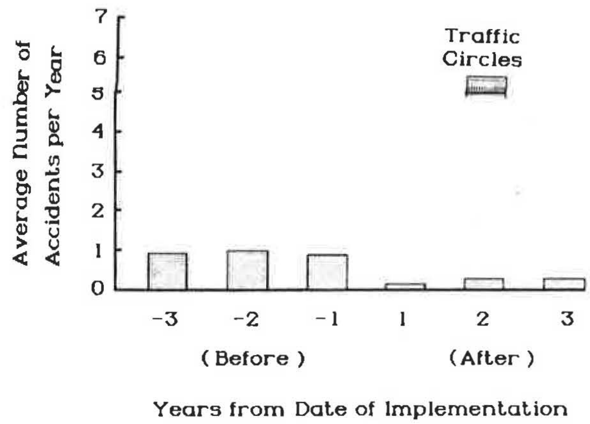


FIGURE 4 Before-and-after accidents at locations using traffic circles.

are the yearly average for four intersection approaches):

	No. of Accidents		Reduction
	Before	After	(%)
Yield	0.70	0.63	10
Circle	1.03	0.63	39

All reductions are significant at $\alpha = 0.05$.

Speeds

In Seattle there is little if any information collected about speeds when yield signs are installed at residential intersections. The literature search done for this study did not reveal any studies conducted with yield signs to determine effects on midblock speeds. The FHWA study done in 1981 concluded that the yield-controlled intersections produced the shortest travel times through an intersection when compared with stop-controlled and uncontrolled locations (3). Without collecting and analyzing additional before-and-after speed data it would be difficult to state that yield signs could significantly decrease midblock speeds.

Seattle has some documentation on before-and-after speeds near intersections with traffic circles. A search through files produced a sample of 10 locations that had speed studies with large numbers of cars and with the same location used for both the before-and-after speed surveys.

At 9 out of 10 traffic circle locations studied, there were decreases in midblock speeds. All of these decreases were statistically significant. The before-and-after speeds are shown in Figure 5.

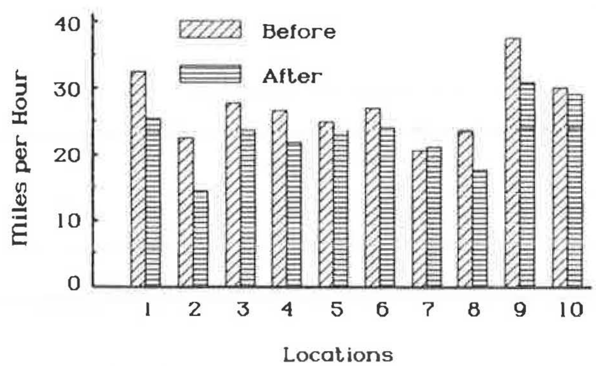


FIGURE 5 Before-and-after midblock speeds at 10 traffic circle locations.

Volumes

Volume information is collected for each location being considered for a traffic circle. For this study, a group of 20 traffic circle locations was used to determine the changes in volumes. These circles were installed in 1983 and had complete records for both before-and-after volumes. There was no before-and-after volume information collected for the yield sign locations studied.

Numbers used for comparisons represent the AWDT volume for the major street. The counts were adjusted for monthly differences.

Of the 20 locations analyzed, 9 locations had increases and 11 locations had decreases after traffic circles were installed. The group as a whole had a decrease of 2 percent in the total volume.

Costs

Another difference about the use of traffic circles is that these devices cost much more than two signs. Construction costs for a traffic circle with landscaping are $3,400; total costs including planning and engineering are $5,550. A pair of yield signs or stop signs costs about $500. It would not be appropriate to compare the cost of installing a traffic circle at a particular location with the cost of installing yield signs or stop signs at the same location. It is obvious that signing would be the less expensive alternative. However, if one values the reduction of the number of traffic control signs in residential areas, traffic circles may be worth the added expense. In some cases people in the neighborhoods have contributed funds to install traffic circles.

CONCLUSIONS

Conclusions drawn from the data analysis and literature review are as follows:

1. Accident reductions: (a) traffic circles are shown to have accident reductions of 77 percent when comparing 3-year before-and-after totals of intersection accidents, and (b) yield signs also have shown a reduction of 77 percent.
2. Midblock speeds: Traffic circles tend to significantly decrease the speeds of vehicles downstream of intersections with circles.
3. Volumes: Locations with traffic circles tend to show a wide range of volume increases and decreases, but the locations studied in this paper had an overall decrease of 2 percent, which is insignificant.
4. Costs: Installation of a traffic circle at a residential intersection costs much more than installation of either a pair of stop signs or a pair of yield signs.

REFERENCES

1. All-Way Stop Study. Office for Planning, Seattle Engineering Department, Seattle, Wash., March 1983.
2. A Study of the Use of Stop Sign Warrants. Technical Committee Report. ITE, San Francisco Bay Area Section, San Francisco, July 1983.
3. W.R. Stockton, R.Q. Brackett, and J.M. Mounce. STOP, YIELD, and No Control at Intersections. Report FHWA/RD-81/084. FHWA, U.S. Department of Transportation, June 1981.

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

Preferential Control Warrants of Light Rail Transit Movements

A. ESSAM RADWAN and KUO-PING HWANG

ABSTRACT

The goal of this paper is to demonstrate a method for evaluating a preferential treatment technique for light rail transit (LRT) in urban areas. A mathematical delay model, which uses probability expressions, is presented to evaluate two LRT preemption signal strategies in existing arterial medians. The model permits the user to evaluate three operational options: a two-phase signal plan, a three-phase signal plan with a separate LRT phase, and a three-phase signal plan with an exclusive left-turn phase for main arterial vehicles. The signal controller modeled in this paper has green extension and red truncation capabilities. Model testing and validation proved that the model parameters consistently produced reasonable results. Control warrant guidelines were developed for two operational options.

Light rail transit (LRT) is catching the attention of numerous cities across North America today. New LRT operations were initiated in Edmonton and Calgary, Alberta, Canada, in 1978 and 1981, respectively (1,2). New systems are in an advanced stage of construction in Buffalo, and others are being considered for upgrading in Pittsburgh, San Diego, and San Francisco.

LRT, as defined by the Transportation Research Board Committee on Light Rail Transit, is a mode of urban transportation that uses predominantly reserved but not necessarily grade-separated rights-of-way. Electrically propelled rail vehicles operate singly or in trains. Most of the LRT operating environment are at grade but with predominantly controlled rights-of-way. Separated right-of-way, on-street operation, and transit-pedestrian malls are the most common forms of at-grade operating environments. Median LRT treatment is a special design in which the light rail line is accommodated in an existing wide median of a multilane arterial. Such design may occur for a heavily traveled arterial, in which case the signal timings should be carefully studied to maximize system passenger throughput. A common preferential control technique for LRT is to use traffic signal preemption in favor of the LRT; however, this technique may adversely affect overall system performance. The major objectives of this study are to investigate preferential control of LRT by using different signal preemption strategies and to attempt to develop control warrants for these strategies.

BACKGROUND INFORMATION

The use of unconditional traffic signal preemption generally results in some loss in intersection capacity. This loss is proportional to the LRT frequency and the particular preemption strategy used. In a recent study (3) the impact of signal preemption on intersection capacity was evaluated. It was concluded that, at a standard intersection at which all other traffic must stop to allow the LRT vehicle to pass, around 10 percent of the available signal time would be lost if preemption occurred every 3 min. Furthermore, for a multilane arterial with far-side transit stops and a constant main-street traffic volume of 20,000 vehicles per day and a cross-street volume range of 10,000 to 20,000 vehicles per day, it was found that a multiphase traffic signal makes LRT preemption feasible in every third cycle. If simple two-phase signals are used and left turns are prohibited, LRT preemption in every second cycle is feasible. Similar capacity analyses performed for a midblock crossing of a four-lane arterial indicated that preemption is feasible as often as every 2 min for traffic volumes as high as 25,000 vehicles per day.

In another study (4) the use of level-of-service criterion to evaluate LRT impacts on traffic flow over arterials was criticized because it significantly favors the automobile mode over the LRT mode and it does not consider the volume of people carried by transit. A factor that indicates the percentage of theoretical capacity of the intersection that is being used (intersection utilization factor) was used to evaluate the impact of operating LRT within the same vehicular right-of-way on street traffic performance. Utilization factors were calculated for three alternative operational strategies:

1. Left turns from the arterial onto the cross street (across the LRT tracks) controlled with a special signal phase,
2. Left turns prohibited from the arterial onto the cross street, and
3. All traffic stopped during LRT passage.

The utilization factors without LRT preemption were also included for comparison. Analysis of these results pointed out a key conceptual difficulty with the use of the traditional level-of-service approach. The results imply that, as the frequency of the LRT operation increases, the feasibility of preemption decreases; it causes an "unacceptable" impact on cross traffic. However, higher-frequency LRT operation actually may mean that greater numbers of transit passengers are traversing the intersection. Thus the true situation may be the opposite from the situation implied by the utilization factor results.

A parametric analysis was conducted in the same study, using a delay model developed by May and Pratt (4), to alleviate the problems with the level-of-service approach. Two major conclusions were drawn: First, the justification for priority treatment for LRT generally increases as the line volume increases, until the headways are so short and cross-street volumes are so high that they begin to greatly increase automobile delay. Second, it was found that preemption can be justified for a large number of LRT headways and combinations of cross-street volumes, whereas the utilization factor criterion resulted in many more design combinations falling into the so-called unacceptable category. Other studies (5,6) involved the development of two macroscopic delay models for the purpose of evaluating the impact of bus signal preemption on street vehicular delay.

The literature review revealed that previous studies have used simple delay models with no capability of evaluating different preemption strategies (green extension and red truncation) and, more important, they all failed to define general warrant guidelines for using signal preemption in association with LRT traffic.

RESEARCH OBJECTIVES

The major objectives of this research study are to develop a mathematical model that estimates private automobile and LRT delays for signalized intersections operating under preemption scenarios, to apply the model to three operational strategies and check its validity, and finally to use the model to develop warrants for signal preemption of LRT movements.

DELAY MODEL

A modified version of Webster's delay model was selected for this research (7); the average delay per vehicle is determined from

$$d = 9/10\ \{[c(1 - \lambda)^2/2(1 - \lambda x)] + [x^2/2q(1 - x)]\}\quad (1)$$

where

- d = average delay per vehicle on the particular intersection approach,
- c = cycle time,
- λ = proportion of the cycle that is effectively green for the phase under consideration (g/c),
- q = flow,
- s = saturation flow, and
- x = degree of saturation.

Equation 1 was used to estimate the average delay per private automobile and LRT. The probability of signal preemption was estimated for each LRT detection event. Signal cycle length and corresponding phase splits were also determined for each detection scenario. The average delay per vehicle and the probabilities were combined, and the estimated delay for preemption and nonpreemption cases were calculated and compared.

Model Assumptions

The following assumptions were made to formulate the analytical model:

1. Pretimed signal controller with a two- or three-phase plan and a cycle length are determined from Webster's optimum cycle formula (7);
2. Minimum red phase durations for main and cross streets are determined from Webster's minimum cycle formula;
3. Absolute minimum cycle length is 40 sec for two-phase and 50 sec for three-phase plans, and absolute maximum cycle length is 120 sec for two-phase plans and 150 sec for three-phase plans;
4. Minimum green phase duration is 12 sec for through maneuvers and 15 sec for left-turn maneuvers;
5. Left-turn adjustment factor is 1.75 for private automobiles; and
6. LRT arrivals follow a discrete uniform distribution or a Poisson distribution (the model was formulated in a manner to give the user the option of using either distribution).

Pedestrian movement can adversely affect the signal preemption process. If the cross-street green phase is constrained by pedestrian clearance considerations, red truncation may not be feasible and the minimum green-phase duration threshold (12 sec) has to be increased. This study did not include the impact of pedestrian movement on LRT priority schemes; however, the model can be adjusted to take into account those impacts.

Probability Expressions

Probability expressions for LRT arrivals during different time periods of the signal cycle were derived for three signal timing strategies. The first strategy (Option 0) is a two-phase plan with prohibition of left-turn maneuvers from the major arterial to the side street; the second strategy (Option 1) is a three-phase plan in which an exclusive phase is dedicated to LRT movements of 15 sec duration; and the third strategy (Option 2) is a three-phase plan in which an exclusive left-turn phase is provided for automobile traffic to turn from the major arterial to the side street. The signal phase durations are shown in Figure 1 and the probability expressions for a selected option (Option 0) are given in Table 1. The detailed derivation of the five probability expressions is beyond the scope of this paper. The probability expressions of Options 1 and 2, and the mathematical derivations, can be obtained from the authors.

MODEL TESTING AND VALIDATION

The probability expressions and the delay equations were coded into a computer program to facilitate and speed up the calculation of delays. The program calculates internally the total delay of private automobiles and LRT under both preemption and nonpreemption strategies and provides the total delay saving (or losses) caused by the preemption. The major input parameters to the model are as follows:

1. Major arterial volume (private automobiles),
2. Cross-street volume (private automobiles),
3. LRT volume per hour,
4. Private automobile occupancy (passengers),
5. LRT occupancy (passengers),
6. Saturation flow rates for major and cross streets, and
7. Advance detection period (sec).

The output measures of effectiveness are as follows:

1. Main arterial nonpreemption delay (private automobile and LRT),
2. Cross-street nonpreemption delay (private automobile and LRT),
3. Main-street preemption delay (private automobile and LRT),
4. Cross-street preemption delay (private automobile and LRT), and
5. Total intersection saving (or losses).

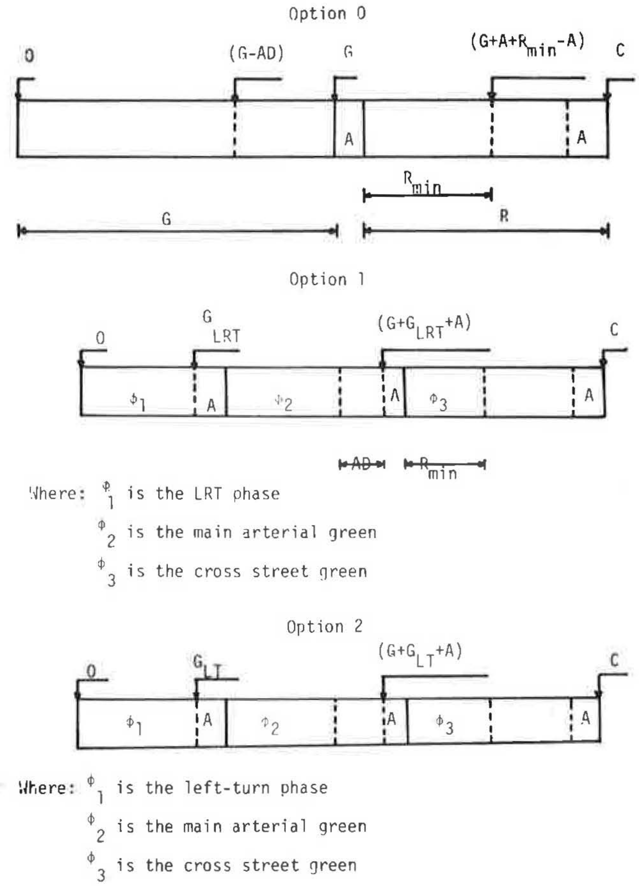


FIGURE 1 Signal timing components for the three options.

A series of runs was conducted to evaluate the model consistency and validity. First, for Option 0, the model was tested for four variations:

1. Variations in main-arterial private automobile volume,
2. Variations in cross-street private automobile volume,
3. Variations in LRT volume, and
4. Variations in advance detection duration.

The results of these runs are shown in Figures 2-5. Figure 2 shows parabolic-like shaped relationships between the main arterial volume and the total intersection gain. One plot corresponds to the uniform distribution of LRT arrivals and the second corresponds to the Poisson distribution. As the plots show, little difference between the two distributions is observed; therefore, it was decided to use the Poisson distribution for the remaining plots only for demonstration purposes. An opposite parabolic shape was observed between cross-street volume and the total intersection gain as shown in Figure 3. The impact of LRT volume on the total intersection gain was observed to be directly linear as depicted by Figure 4, and little variation was noticed between the eight levels of advance detection period as shown in Figure 5.

A second testing was conducted for Option 1 in which a three-phase signal plan was valuated with a dedicated phase for LRT traffic. Different levels of main arterial volume were tested, and the results are shown in Figure 6. It was concluded that the addition of an exclusive LRT phase adversely affects the total intersection gain. As for Option 2, a fixed left-turn volume was assumed at 100 cars per hour, and different levels of main arterial volume were evaluated. Figure 7 shows the results of Option 2 testing, in which a sharp decline in the total intersection gain with the increase in main arterial volume is observed.

PREFERENTIAL CONTROL WARRANTS

The model was applied to a wide range of traffic volumes on main arterials, cross streets, and LRT for Option 0, and a regression analysis was attempted to correlate these variables with the total intersection gain. The following model was attained:

$$\text{Gain (passenger-sec)} = -30481.75 + 1742.70\ \text{LRT} - 61.68\ \text{PC1} + 117.70\ \text{PC2} \quad (R = 0.88) \tag{2}$$

where

PC1 = main-arterial volume (cars/hr),
PC2 = cross-street volume (cars/hr), and
LRT = light rail transit volume (trains/hr).

The signs of the independent variables agree with previous findings, and the regression equation was used to develop signal preemption warrants under different demand levels. By substituting zero in Equation 2 and using PC1 constant values of 400, 600, and 800, boundary lines of the control warrant regions were developed (see Figure 8).

As for Option 1, it was found earlier that no gain can be realized under any demand levels and therefore no attempt was made to develop warrant

TABLE 1 Probability Expressions for Option 0

No.	Event	Discrete Uniform Distribution	Poisson Distribution
1	No LRT arrival during a cycle	M = (c) (LRT)/3,600 If M < 1, P_1 = 1 – M If M ≥ 1, P_1 = 0, M = 1	P_1 = EXP(-LRT • C/3,600)
2	LRT arrives in a cycle and no preemption occurs	P_2 = (G + A – AD) (M)/C	P_2 = EXP[-(LRT) (C – A – G + AD/3,600)] – EXP[-(LRT · C/3,600)]
3	LRT arrives during a cycle and there is red truncation	P_3 = (Rmin) (M)/c	P_3 = EXP(-LRT · AD/3,600) – EXP[-(LRT) (AD + Rmin)/3,600]
4	LRT arrives during a cycle such that red truncation occurs after Rmin	P_4 = (C – A – G – Rmin) (M)/C	P_4 = EXP[(-LRT) (Rmin + AD)/3,600] –EXP[-(LRT) (C – A – G + AD)/3,600]
5	LRT arrives during a cycle such that a green extension occurs	P_5 = (AD) (M)/C	P_5 = 1 – EXP(LRT AD/3,600)

Note: LRT = light rail transit flow, C = cycle length, G = main arterial green period, A = amber phase duration, AD = advance detection period, and Rmin = red phase due to red truncation.

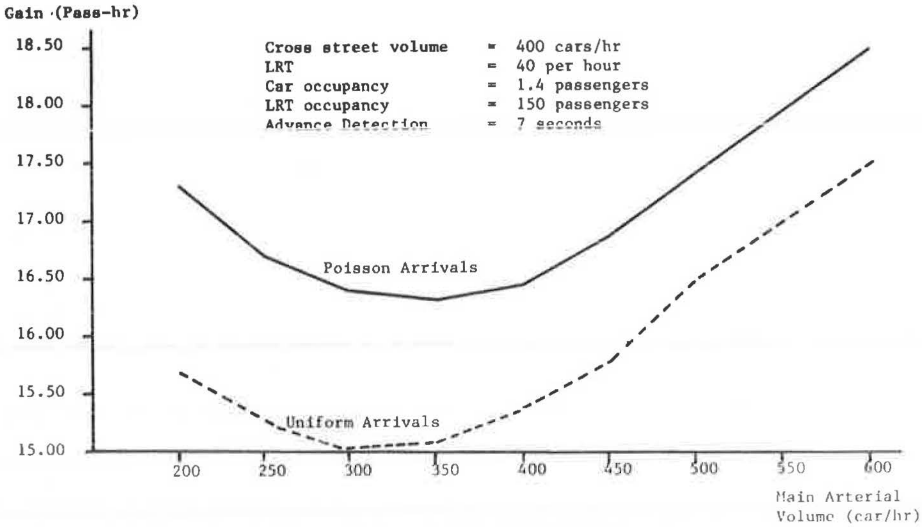


FIGURE 2 Passenger delay gains due to variations in main arterial volume.

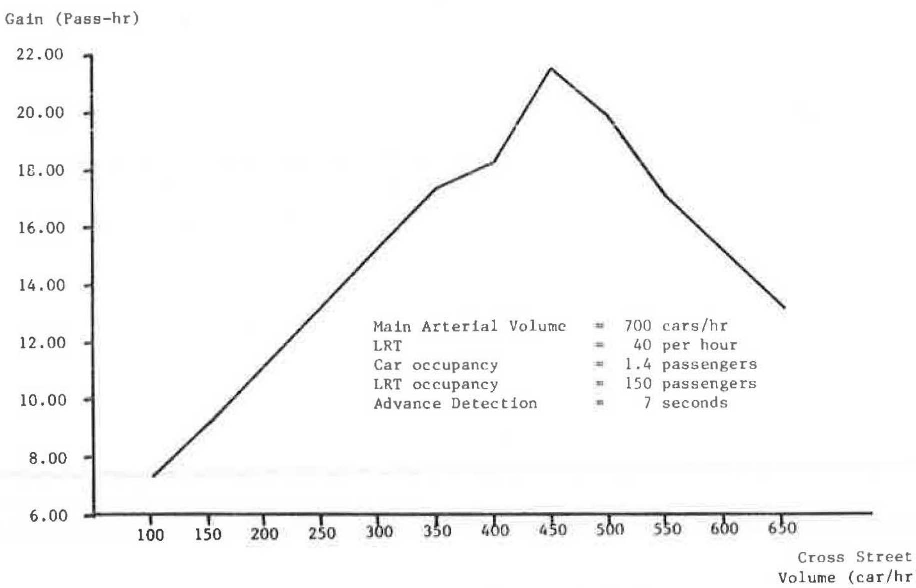


FIGURE 3 Passenger delay gains due to variations in cross-street volume.

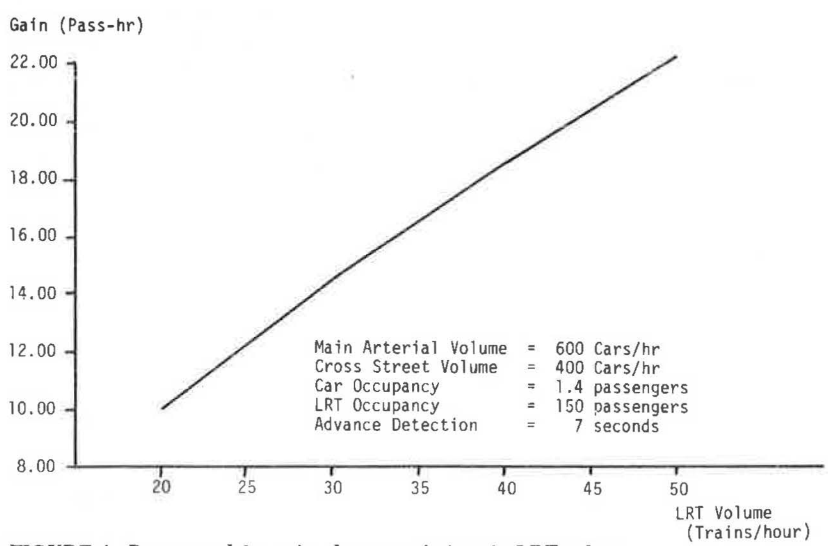


FIGURE 4 Passenger delay gains due to variations in LRT volume.

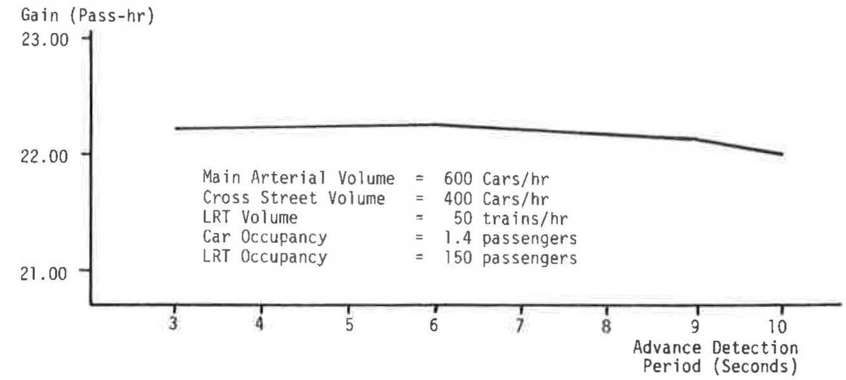


FIGURE 5 Passenger delay gains due to variations in advance detection durations.

regions. For Option 2, the process was repeated, and a regression model was calculated:

$$\text{Gain} = 1163.80 - 34.79\ \text{LRT} + 2878.2\ \text{PLT} + 2.15\ \text{PC2} \quad (3)$$

$$(R = 0.904)$$

where

LRT = light rail transit volume (trains/hr),
PLT = percent left turn, and
PC2 = cross-street volume (cars/hr).

The negative sign of LRT is expected because as the LRT volume increases the total LRT passenger delay increases during the exclusive left-turn phase, and consequently the overall intersection gain decreases. On the other hand, as the percentage of left turns increases, more left-turn traffic uses the third phase and the overall intersection gain increases. The control warrant regions for this option are shown in Figure 9.

SUMMARY AND CONCLUSIONS

The purpose of this paper was to demonstrate a method for evaluating and testing signal preemption strategies of LRT movements in existing arterial medians. Three operational options were identified,

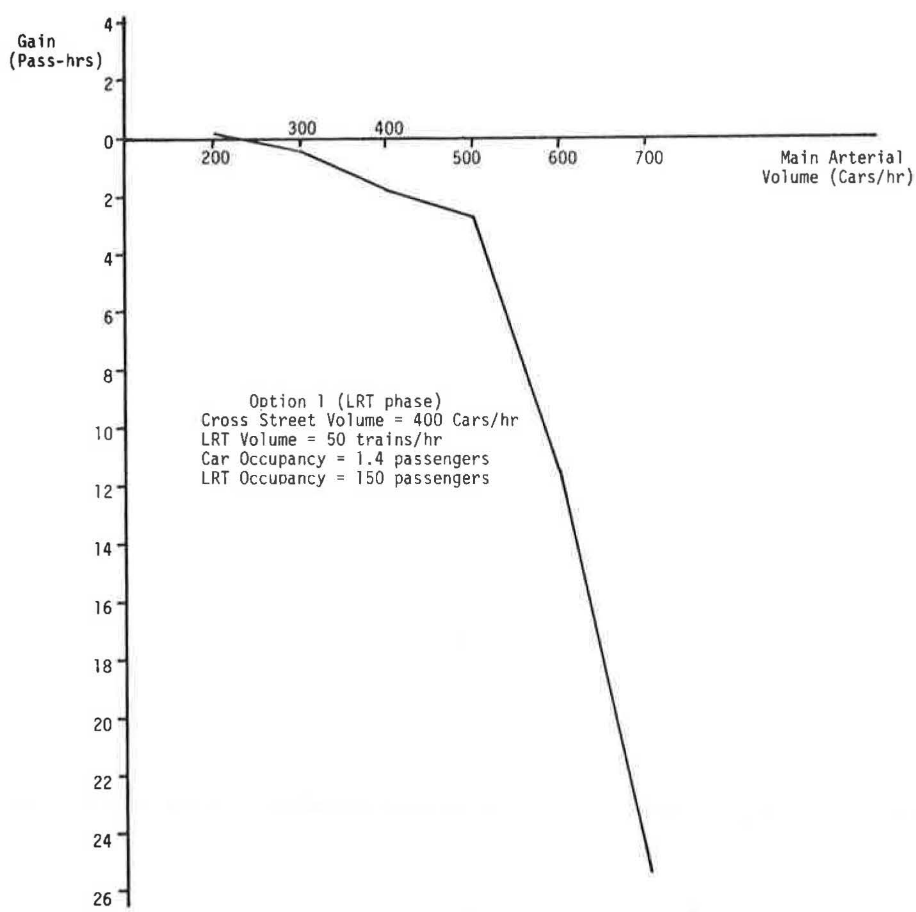


FIGURE 6 Passenger delay gains due to variations in main arterial volume for Option 1.

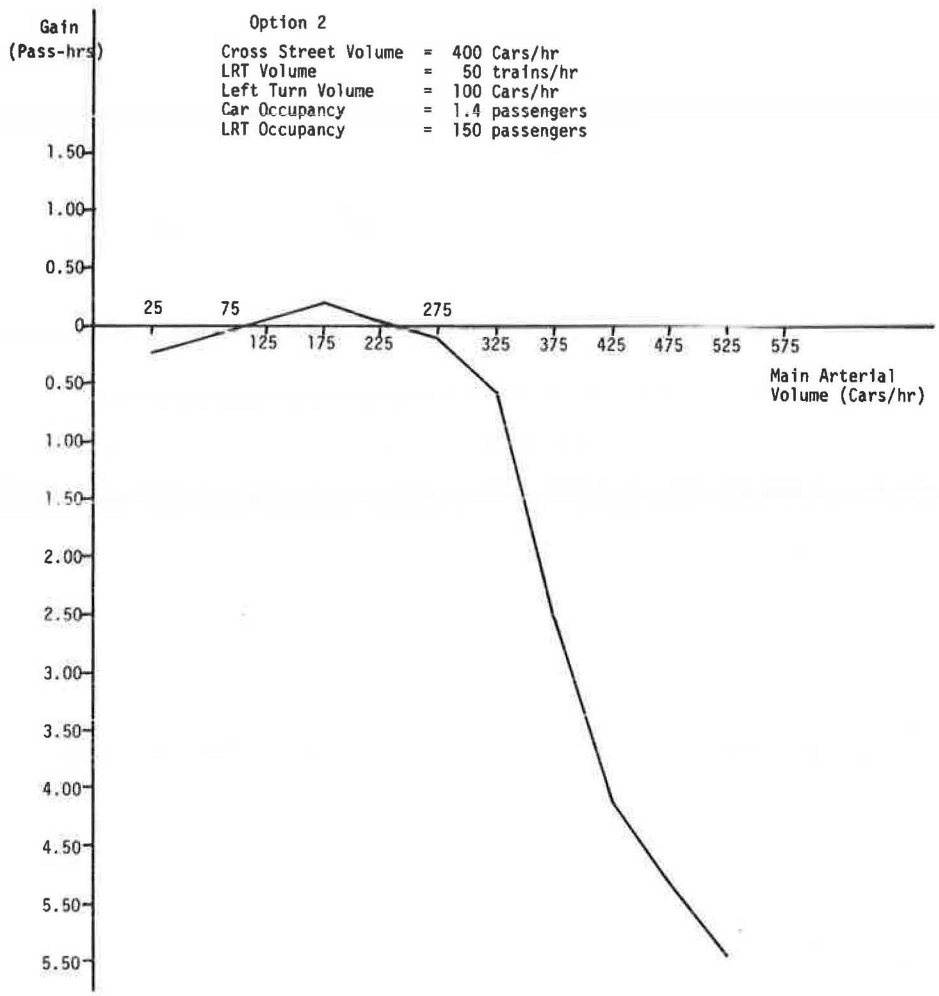


FIGURE 7 Passenger delay gains due to variations in main arterial volume for Option 2.

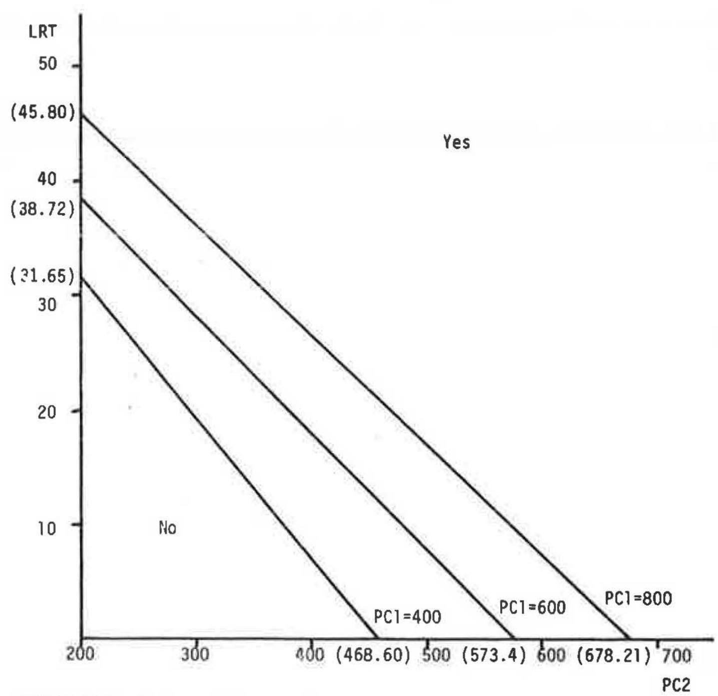


FIGURE 8 Preferential control warrants for Option 0.

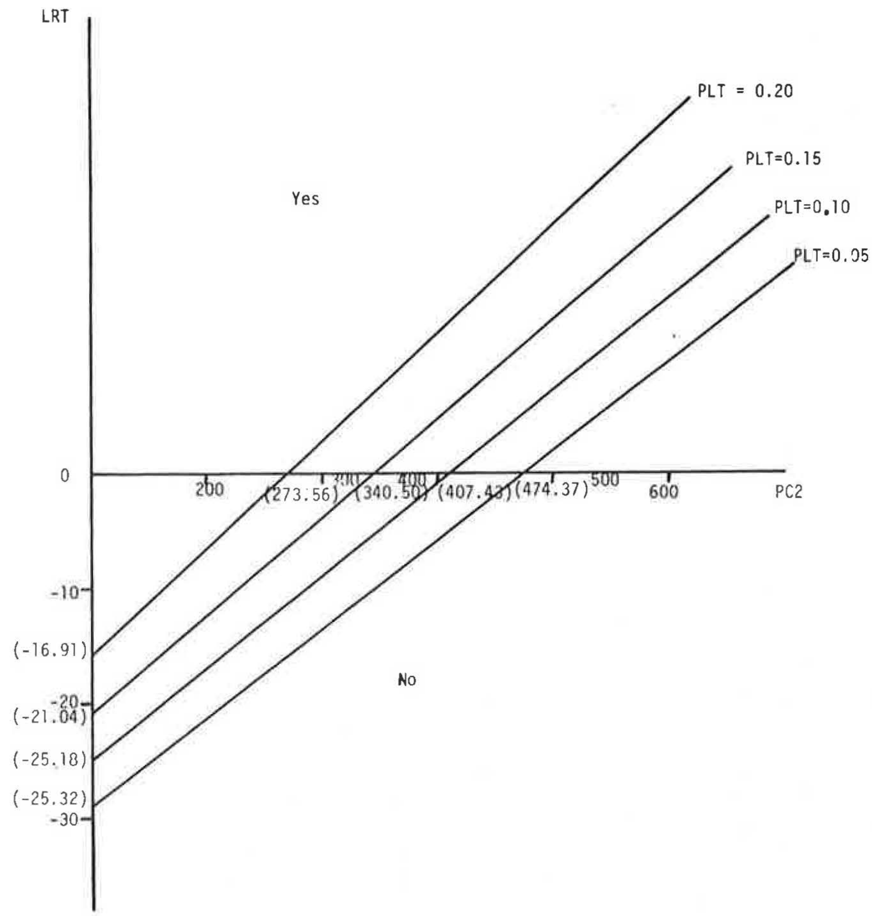


FIGURE 9 Preferential control warrants for Option 2.

and the probability expressions for a selected option were documented. Webster's delay model was adopted to estimate the average delay per vehicle per approach.

The model was tested by using a set of hypothetical demand parameters to validate the model. The results of the model testing proved that the model parameters consistently produce reasonable results, and that the model is sensitive to variations in the main arterial and cross-street volumes. Furthermore, it was concluded that for the two-phase signal plan (Option 0), the overall intersection gain due to signal preemption is linearly proportioned to LRT volume, and that there was no impact of advance detection duration on the intersection gain. It was also found that for the three-phase signal plan with a separate LRT phase (option 1), no intersection gain was observed for almost all main arterial volume levels. As for the three-phase signal with an exclusive left-turn phase (Option 2), it was found that there exists an optimum main arterial volume at which the overall intersection gain is maximum for a given constant left-turn volume. Finally, boundary lines of the control warrant regions for Options 0 and 2 were developed in a chart format.

REFERENCES

1. Technical Council Committee 5C-5. Light Rail Transit in North America: What's Going On? ITE Journal, March 1980.
2. E. Guillot. LRT Design Choices: Edmonton and Calgary. Transportation Quarterly, July 1983.
3. H.W. Korve. Traffic Engineering for Light Rail Transit. In TRB Special Report 182: Light Rail Transit: Planning and Technology, TRB, National Research Council, Washington, D.C., 1978, pp. 107-115.
4. T.J. Stone and W.A. Wild. Design Consideration for LRT in Existing Medians: Developing Warrants for Priority Treatments. In TRB Special Report 195: Light Rail Transit: Planning, Design, and Implementation, TRB, National Research Council, Washington, D.C., 1982, pp. 170-175.
5. A.E. Radwan and J.W. Hurley, Jr. Analytical Model of Bus Preemption Signal System: A Macroscopic Approach. Transportation Quarterly, Jan. 1982.
6. A.E. Radwan and J.W. Hurley, Jr. Macroscopic Traffic Delay Model of Bus Signal Preemption. In Transportation Research Record 881, TRB, National Research Council, Washington, D.C., 1982, pp. 59-65.
7. F. Webster. Traffic Signal Settings. Technical Paper 39. Road Research Laboratory, London, England, 1958.

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

Evaluation and Improvement of Inductive Loop Traffic Detectors

EDWARD W. BIKOWITZ and SCOTT P. ROSS

ABSTRACT

It was determined in 1980 that approximately a quarter of New York State's 15,000 inductive loop detectors, used to control traffic signals, were out of order at any given time and were maintenance-free for an average of only 2 years. A study was made to find the major causes of loop failures and how to reduce them. Installation methods in New York and elsewhere were investigated, and hundreds of failed loops were studied to find failure types and causes. Data suggested that failure was mainly caused by improper installation, inadequate loop sealants, or wire failure. Encased wire (regular signal wire protected by continuous smooth-bore polyethylene tubing), which provided greater freedom of movement in areas of pavement distress, was being evaluated in other states. Also, instead of diagonal cuts at the corners of the loop, it was decided to cut the corners at 90-degree angles and then chisel or core them, thus saving sawing time and equipment wear and also having negligible impact on the loop wire itself. Another cause of failure is damaged or broken wire because of its floating to the surface of the sealed slot, which can be avoided with a simple hold-down device. Other recommendations included saws with greater horsepower, and complete pressurized washing and drying of the saw slot to enhance sealant bond. Laboratory tests were developed to evaluate sealants before purchase to assure that those used would provide strength, longevity, water resistance, good bond to the pavement, flexibility, wire encapsulation, and ease of installation. New York's new methods and materials appear to be the best currently available, and a program has been established for continued evaluation of these detector systems. A special training video tape is now available that covers these new materials and procedures.

Traffic loops are an integral part of today's highway system; they are used to control traffic flow at intersections by providing information to microprocessors that control signal patterns. Current estimates place the number of loop detectors maintained by the New York State Department of Transportation (NYSDOT) at about 15,000--a number that is increasing as improved traffic control is sought. The current trend in New York is to couple a loop detector with a computer to improve traffic control. The computer-enhanced system, although more costly, has the added ability of adjusting signal timing to meet traffic demands during different time periods throughout an entire area, rather than just one intersection for a fixed time period. With increased sophistication, these systems are becoming ever more dependent on continued successful operation of in-pavement loop detectors.

The inductive loop detector (ILD) consists of a specified number of turns of wire buried in the pavement and connected to a detector unit or amplifier. Current is passed through the wire loops to create a magnetic field; as vehicles break the field the overall inductance of the loop circuit changes. The detector senses this change and sends a message to the controller circuit, which adjusts the light signal according to its programming. Unfortunately, reliability of these loops has not been good, resulting in expensive replacement and, perhaps more important, in serious delays and inconvenience to motorists. In the past most inductive loops were maintenance free for about 2 years. By then systems began to suffer about 25 percent failures, thus requiring loop repair. The investigation reported here was initiated to study inductive loops and find ways to increase their reliability and lengthen service life.

Various means were employed to determine how to increase loop life: field evaluation of current failures, questioning of maintenance crews, a literature review, and questionnaires to other states. Results from these inquiries could be classified into two areas: materials and installation techniques. Each required extensive changes to correct their deficiencies.

INSTALLATION TECHNIQUES AND IMPROVEMENTS

Reliability Problems with Installation

Before discussing physical changes in installation methods, it should be pointed out that one problem area--lack of consistency in loop installation from region to region--was solved. Signal-loop materials and installation techniques are controlled by state standards and specifications. New York installs traffic loops either by contract or by department maintenance forces, and although all persons should follow the specifications, installers frequently varied slightly in methods and materials. As a result of this study, stricter policy was established to ensure that different areas used the same techniques and materials. This policy was explained as an attempt to establish known controls against which to compare experimental techniques rather than as an attempt to limit their efforts to find better procedures.

The investigation revealed several areas within the installation process that needed revision to improve loop reliability and several other areas in which speed, safety, and efficiency could be improved. The areas involving reliability were corner cutting, wire floating, splicing at the pavement edge, and cleaning of the sawed slot. Areas involving efficiency and safety were type of saw blade to be used, saw power [9.2 versus 18.09 horsepower (hp)], dry versus wet cutting, and use of small air tools.

One of the main causes of loop failure was wire breakage. In terms of installation problems, this could be related to several areas: sharp corners in the slots wearing through the loop wire, the loop floating to the top of the slot and being exposed to traffic wear, splicing of the loop wire and detector leads at the pavement edge, and failure of the loop sealant to bond to a dirty or wet slot.

Before this study, the loop corners were cut on diagonals (Figure 1). This did reduce some of the corner-cutting problem, but created another problem in that pie-shaped segments of pavement could break out, thus exposing whole corners of the loop to traffic and weather. The solution chosen was to saw the loop in four straight cuts, chisel out the corners, and round them off as smoothly as possible.

Once a slot is cut, it must be cleaned before installing the wire. If not properly done, many fine particles remain in the slot, thereby decreasing the chance of a good bond between the sealer and the sidewalls. A poor bond eventually causes the sealer

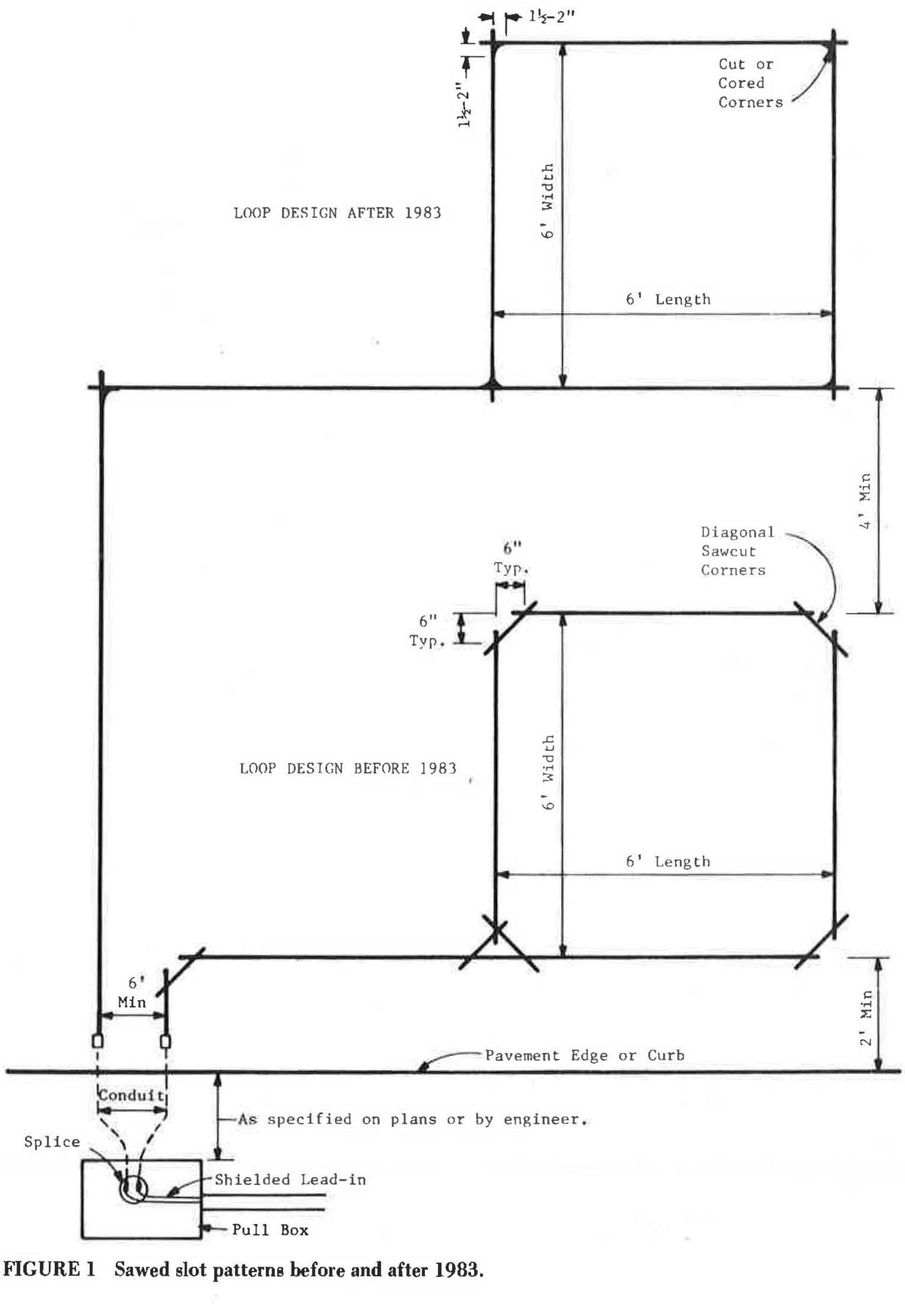


FIGURE 1 Sawed slot patterns before and after 1983.

and wire to pop out of the loop, which results in early exposure to traffic and weather. The former method, using only the air supply from a large compressor, was inefficient. A pressurized water system was desired, but the increased cost of having a water-pumping truck was not justifiable. Instead, a special nozzle was devised to combine a standing gravity-fed water supply and the already available air compressor. Using the Venturi principle, a nozzle was fabricated to supply pressurized water. The air supply from the compressor passes through a restricted chamber to propel the water from the standing water supply. This system resulted in a much cleaner slot. The same nozzle is used to dry out the slot by simply shutting off the water supply. If the compressor were to have hot-air capability, the work would proceed even faster. It should be noted that once this nozzle was shown around the state, many department maintenance forces found uses for it in other areas.

At this point in the installation process the wire is placed in the loop starting and ending at the pavement edge. Wire continuity is checked now, and if there is no problem the loop wire is spliced to the detector leads. As in any electrical installation, the splice is one of the most critical components of the electrical circuit. Past experience indicated that the method of splicing varied from region to region, and no one method was more reliable than any other method. Evaluation of all current techniques plus some questioning of the research electronics staff resulted in the following splicing specification.

An uncoated, metal, solderless crimp connector or solder or both are used to make the initial connection. After joining the wires, liquid waterproofing is applied. Next, heat-shrink tape or rubber tape is applied over the splice, making sure the outside tubing is securely joined to the encased wire. The advantage of using heat-shrink tape is that there is less seam area for possible debonding. After applying the heat-shrink or rubber tape, another coat of waterproofing material is recommended. Next, layers of polyvinyl chloride (PVC) tape are added in combination with waterproofing material. A final layer of PVC tape is applied, followed by a waterproof coat. A third layer of heat-shrink or rubber tape is added, followed by another waterproof coating. The signal wire should be reinsulated with the proper combination of materials to equal 1.5 times the original wire insulation thickness.

At this point the encased wire must be secured in the slot. It must be prevented from floating to the top of the sealant during curing or during hot summer days when the sealant softens. A hold-down was required, but the type of material to be used had never been specified. The study indicated that most hold-downs failed to perform properly; that is, some absorbed water, melted at higher sealant curing temperatures, lacked holding ability, lacked recovery from deformation, or were difficult to install. One material found to perform well was an open-celled backer rod. It is readily available, inexpensive, and easy to use; compresses well; does not wick water; and some brands resist intermittent temperatures in excess of 400° F. One-inch strips were installed every 2 ft, taking care to include strips on either side of the corners or wherever the wire changed direction.

The sealant, when properly mixed and prepared, is poured into the slot. In the past it was found that not enough emphasis was placed on leveling the sealer. If the sealer is higher than the pavement surface, it is exposed to both tire and snowplow wear. A snowplow or even a car conceivably could hit the protruding sealant and jerk it and the wire from the slot. To handle this, department forces developed a special tool: a V-shaped rubber-bladed squeegee on a 4-ft wooden handle, made from old broom handles and old rubber mudflaps. This tool proved adequate and convenient for pulling sealant into the slot, leveling it, and also reducing waste. At this point the sealant surface, once cured, may be dusted with cement dust or dry silica sand to eliminate any slight surface tackiness, and traffic is allowed to resume as quickly as possible. Road dust, gravel, or stone dust should not be used because incompressibles may be introduced into the loop, thus causing problems later.

Speeding Up Installation

From the field study and discussions with maintenance personnel, it was decided that the speed of installation could be increased, thereby reducing costs. The major area in which operations could be accelerated was cutting of the slot. Maintenance forces believed that the 9.2 hp of the saw then supplied was too low for the heavy-duty cutting required. This could be justified because of the change in the wire diameter called for in the new specifications (i.e., larger-diameter wire required a larger-width slot). Advice on horsepower was obtained from several sources, including regional maintenance crews, contractors, rental agencies, saw manufacturers, and other states. Almost all agreed that saws were underpowered even when cutting 1/4-in.-wide slots, and to try to cut 3/8-in. slots with the same saw would result in much slower work and more saw breakdowns. Most agreed that the minimum power of the saw should be 18 hp. The regions are now receiving 18-hp saws as replacements for 9.2-hp saws.

Along with the increase in the power of the saw, changes were also made in the type of blade used. Regions varied as to whether abrasive or diamond blades should be used. Also, by varying the blade type a change is made from using a wet-cutting method (diamond) to using a dry-cutting method (abrasives). Dry cutting creates large dust clouds that are irritating and dangerous to the work crews, passing motorists, and nearby neighbors. Wet cutting with a diamond blade produces fewer hazardous side effects. The only advantage of the abrasive blade over the diamond blade is initial cost, but when life expectancies of the two kinds of blades are compared the diamond blade wins by a margin large enough to overcome the cost difference. The life expectancy of diamond blades is equal to 35 abrasive blades, and the diamond blade could reduce cutting time by two-thirds.

It was also necessary to select the proper type of diamond blade. Diamond blades are rated to cut specific types of pavement. In New York it is estimated that 75 percent of all state-maintained signal loops are located in some type of asphalt road surface. A diamond blade rated to cut asphalt was selected for test installation and later recommended for statewide use.

Along with the change in the saw and blade, other tool improvements were recommended. Because all maintenance crews have a compressor as part of their assigned equipment when doing loop installations, why not supply them with air-powered hand tools? Chipping the loop corners, finding and removing old feeder tubes, and general slot cleanup could be greatly speeded up by their use. In fact, chipping out by hand with a cold chisel is now the way to install feeder tubes and the extra-wide slot where the loop wire crosses from one slab to another.

LOOP MATERIAL CHANGES

Wire Specification Changes

Field observations, a literature search, and letter surveys all indicated that wire breakage was a major factor in loop failures. New York's current wire standard called for a seven-strand copper UL Type XHHW, No. 14 AWG single-conductor cable, insulated with a polyethylene cover that had an outside diameter of about 0.14 in. Installers believed that this wire was not as durable as others. This wire was totally restrained, and when a pavement moved the wire could be stressed or broken or both. In fact, one department agency is using a solid-core heavy-jacketed wire that does a good job. With this in mind, a search was made to find the best wire for the job.

The search found that only Illinois had studied the problem extensively. They found that encasing the wire in a flexible vinyl tube along with other changes resulted in a large reduction in failures. Based on Illinois' experience and New York's research, a No. 14 AWG, Type THWN or THHN stranded copper wire encased in a vinyl tube was chosen for evaluation (Figure 2). Before field testing, the wire was checked in the laboratory for resistance to temperatures experienced during installation. Once it passed this test, it was taken into the field for final evaluation where it has proved effective.

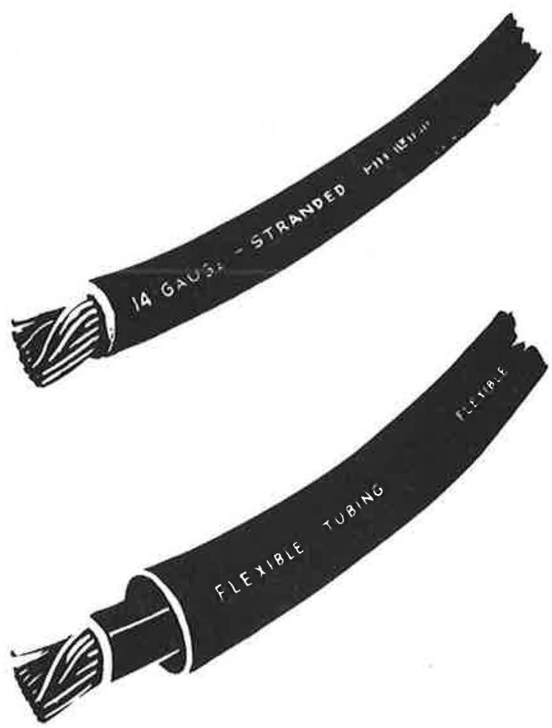


FIGURE 2 Standard 14-gauge stranded wire (top), with polyethylene sheath (bottom).

Sealer Changes

Sealants were known to be a problem from the beginning of this study. New York had no sealant testing procedure, and an informal approved list was based mostly on the manufacturer's sales literature rather than any testing or field results. Field surveys revealed many problems with most sealants in use.

The first types of sealers tested were cold-applied, asphalt-based emulsions. This type of sealant was too thin, ran out on slight grades, and soaked into the subgrade surface. It was also susceptible to washout if there was unforeseen wet weather before it cured.

Caulking tubes of sealant were evaluated and found to be difficult to work with. They were prone to poor encapsulation and had long curing times, considerable shrinkage, and poor bonding. In fact, one important consideration applicable to any sealer is how curing occurs. If it cures by evaporation or loss of volume it is probably not usable because the resulting shrinkage encourages bond failure and eventual loop failure.

Silicon-based material that uses a primer was also tested and rejected. The sealer was too thick and did not appear to encapsulate the wire. It was also found that the primer, a toluene-based material, acted as a cutback on the asphalt and actually weakened the bond area. Field surveys verified that it debonded early.

One of the more popular sealers was hot asphalt. Most states and maintenance crews were found to use this material because of low cost, ready availability, and extensive experience in using it in the past. The hot-asphalt sealer is not recommended for several reasons. First, it is often heated to temperatures exceeding the insulating properties of the wire. Second, it requires frequent resealing from season to season. Third, and most important of all, it is dangerous to work with because it presents not only health hazards because of exposure to fumes but also because of the possibility of fire and explosion while the material is being heated. The dangers involved in its use are enough to disqualify it from consideration as a sealer.

As a result of this study, a laboratory sealant evaluation and testing program was established to set specifications for screening loop sealers before actual highway use. To aid in setting up a specification, the following qualities for a good sealant are desirable:

1. Adequate pot life,
2. Minimal curing time,
3. Sufficient viscosity to encapsulate the wire and not flow out on a slight grade,
4. Sufficient flexibility to accept some pavement movement without cracking the sealant or debonding the wire,
5. Good bond to both concrete and asphalt,
6. Ease of preparation and use,
7. Ease of cleanup,
8. Longevity with a minimum of maintenance,
9. Lack of shrinkage, and
10. Safe use.

The desirable properties were combined with engineering estimates and laboratory results for successful sealants to establish laboratory specifications. Any sealer could be submitted for testing and, if acceptable, would be evaluated in the field for a 6-month test period. This was necessary because laboratory tests could not adequately judge how the sealer would perform in the field. It is hoped that current specifications can be further refined as more information is gathered from laboratory and field evaluations. As a result, New York now has uniform procedures for judging sealers and establishing an approved list.

CONCLUSIONS

As a result of this study, statewide installation methods and specifications have been established.

For three construction seasons, these changes were used in approximately 30 installations, most of which were failed sites. After 3 years, all sites are still operational except for a few damaged by highway reconstruction.

Except for the introduction of encased wire and cold-applied sealants, change in materials is less important than emphasis on correct, uniform methods of signal-wire installation. A survey of the department's regional offices indicated that installation techniques varied among locations because of different interpretations of specifications and availability of materials. Major specification changes were

1. Use of No. 14 AWG stranded, single-conductor wire encased in a continuous vinyl or polyethylene plastic tube;
2. Use of improved roadway loop-embedding sealer; and
3. Use of chipped-out or cored corners instead of diagonal sawcuts at the corners of the loop slot cutouts.

Use of state-specified encased signal wire makes it necessary to saw a 3/8-in.-wide slot. The corner diagonal saw cuts have been replaced with chipped or cored corners. Hot bituminous-based sealants have been replaced by an approved list of cold-applied sealers. However, these changes are less critical to loop longevity and operation than the proper loop installation method, which should be standardized throughout the state.

In an attempt to standardize installation methods, a loop-wire informational seminar was conducted. Representatives from various regional construction and maintenance crews attended a 1-day presentation and a 1-day demonstration of field techniques. A special report was prepared for this seminar (1). This special report and the final report (2) are available from NYSDOT for those seeking more details. A video tape (3) that shows correct installation techniques has been prepared by NYSDOT for FHWA and is now available as a training film.

ACKNOWLEDGMENT

The research reported here was conducted in cooperation with FHWA, U.S. Department of Transportation.

REFERENCES

1. E.W. Bikowitz and S.P. Ross. Installation of Vehicle Detector Loops. Special Report 75. Engineering Research and Development Bureau, New York State Department of Transportation, Albany, July 1983.
2. E.W. Bikowitz and S.P. Ross. Evaluation and Improvement of Inductive Loop Detectors. Research Report 119. Engineering Research and Development Bureau, New York State Department of Transportation, Albany, Feb. 1985.
3. M.A. Flynn. Better Inductive Loop Detectors: The Script of a Training Videotape. Special Report 81. Engineering Research and Development Bureau, New York State Department of Transportation, Albany, Sept. 1984.

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

Evaluation of Reflectorized Sign Sheeting for Nonilluminated Freeway Overhead Guide Signs

M. D. HARMELINK, G. HEMSLEY, D. DUNCAN, R. W. KUHK, and T. TITISHOV

ABSTRACT

A comparative evaluation of various combinations of reflective sheeting (Level 1 = high-intensity, super-engineering grade; and Level 2 = engineering grade) on nonilluminated freeway overhead guide signs was made under road test conditions. A panel of nontechnical observers was used in a subjective evaluation of the signs. Luminance measurements were made by using a telephotometer at the front passenger's eye position using low-beam headlights together with traffic stream headlight illumination. Cost analyses were also performed. The study concluded that, pending the results of further tests on high-intensity versus super-engineering grade for sign message and border, the recommended course of action for freeway overhead guide signs was to implement high-intensity foreground (legend and border) on engineering-grade background.

There has been much recent debate and study of the most cost-effective combinations of external illumination and sign sheeting reflectivity to provide adequate nighttime visibility of freeway overhead guide signs. Sign sheeting is available in three reflectivity levels:

• Reflectivity Level 1 [high-intensity grade (HI)]--highest reflectivity; encapsulated lens reflective sheeting.
• Reflectivity Level 2 [engineering grade (EG)]--lowest reflectivity.
• Super-Engineering grade (SEG)--between Levels 1 and 2 in reflectivity.

In the mid-1970s the Ontario Ministry of Transportation and Communications (MTC) adopted use of both external sign illumination and high-intensity reflective sheeting for freeway overhead guide signs. After a trial period through the winter of 1981-1982, the external sign illumination was discontinued, except in specific critical locations, which resulted in associated savings in energy and maintenance costs. There has been no detectable increase in the number of accidents because of discontinuance of sign illumination, nor have there been public complaints.

The decision to use high-intensity sheeting, applied to both the sign foreground (message and border) and the background, had been based on an expectation of a 15-year life for the material, which, despite its higher capital cost as a single-source product, would result in an equal or better life-cycle cost than engineering-grade sheeting (assumed 7- to 8-year life) and would provide higher reflectivity. Super-engineering-grade sheeting was not available in 1977, but entered the market later as a competitor to high-intensity sheeting.

Recent unit prices of sign sheeting for MTC have been as follows (note that 1 m^2 = 0.0929 ft^2):

	Price ($/$m^2$)	Ratio
HI CDN	39.60	4.00
SEG CDN	23.60	2.83
EG CDN	9.90	1.00

By 1983 high-intensity sheeting durability problems had become apparent and were considerably reducing the effective life of the material. The failure occurred primarily in the large expanses of sign background material. It was decided to test whether more cost-effective combinations of sign sheeting materials could be used for overhead signs without sacrificing nighttime visibility.

It is worth noting here that MTC neither endorses nor condemns commercial products, and it is not the Ministry's intent to do so here. Laboratory tests, field experiments, field experience, and cost comparisons will, from time to time, lead to changes in application decisions. This should not be interpreted as a rejection of a given material or product, but rather as a decision based on cost-effectiveness assessment of a combination of factors at a particular time.

OBJECTIVE AND SCOPE

The objective of the study was to determine the most cost-effective combinations of sign sheeting materials for freeway overhead guide signs without external sign illumination.

Forty-seven signs located in metropolitan Toronto and the surrounding area were used in the study. Twelve were practice signs used to familiarize the subjects with the study procedure, and they were not included in the analyses. Roadway illumination was present throughout the test area, but signs were not illuminated by sign luminaires.

Three different reflective sign sheeting materials were used on the overhead signs. Original signs were constructed of reflective sheeting on extruded aluminium panels. Some of the test signs were refurbished signs that used reflective sheeting on a 1.2-mm (0.040 in.) aluminium overlay that was pop-rivetted to the original aluminium extrusion signs. Except for one sign, no signs were more than 4 years old. The data in Table 1 summarize the relevant parameters of the signs used in the study.

TABLE 1 Parameters of Signs in Study

Material Combination (paired signs)	No. of Signs	Position	Avg Age (months)	Avg Text (no. of letters)	Contrast Ratio[a]
HI/HI	7	Left	29	26	5
HI/HI	9	Right	25	14	5
HI/EG	5	Left	7	19	19
HI/EG	2	Right	8	14.5	19
SEG/SEG	2	Left	20	15.5	3.3
SEG/EG	2	Right	5	12.5	9.5
EG/EG	1	Left	5	12.5	7.4
EG/EG	2	Right	2	16. 5	7.4

[a]As measured in weatherometer tests with an observation angle of 0.2 and an entrance angle of -4 after 1,000 hr of exposure.

The primary comparisons in the test were between signs that use HI/HI and HI/EG sheeting materials. Test signs that use SEG sheeting materials were too few in number to permit more than speculative interpretations.

Most of the test signs were refurbished in the summer of 1983. The sign evaluation was carried out in the spring of 1984, after the signs were exposed to one winter of weathering.

A panel of nontechnical observers was employed in a subjective evaluation of the signs. Luminance measurements were also made by using a telephotometer at a front passenger's eye position and low-beam headlights together with traffic stream headlight illumination.

SUBJECTIVE EVALUATION

Methodology

Nineteen full- and part-time MTC employees took part in the study. None of the observers had any direct involvement with traffic signs in their jobs. They ranged in age from 20 to 60; six were 30 or younger, nine were between 31 and 50, and four were 51 or older. Both males and females took part in the study, and all of the observers were licensed drivers.

The sign evaluation took place on four nights in late April. An evening's evaluation session was cancelled if it was raining at the start of the session. As a result, all of the sessions took place in dry weather. At the beginning of each session the group of observers was given approximately 15 min of instruction about the study procedure. Shortly after it was completely dark (about 7:45 to 8:00 p.m.), the observers began traversing a preset route that required close to 2.5 hr to complete. Two observers rode in a car as front seat passengers with a trained driver. Training for the drivers consisted of learning the test route and practicing it a number of times so that each test sign could be passed at a constant speed of 100 km/h (62 mph) and in the same driving lane. The lane driven in was selected so that the car would pass under the left-most sign of the paired sign(s) to be evaluated. Identical full-sized station wagons were used, and the alignment of their low-beam headlights was set to a standardized specification.

Markers were placed on the roadside 210 m (700 ft) before each sign or pair of signs. The messages on the signs that were evaluated were composed of 38-cm (16-in.) letters. By using the accepted legibility distance of 6 m/cm (50 ft/in.) of letter height, the signs would be legible at approximately 245 m (800 ft). A marker distance of 210 m was used to ensure that observers would be able to read the sign at or near the beginning of the observation period. As each marker was passed, the driver instructed the observers to start observing and forming their judgments of the signs. Therefore, the observation period was more than 7.5 sec. The observers scored the left and right signs immediately after the signs were passed.

The signs were scored on four 7-point scales for brightness, legibility, adequacy, and glare. Brightness (conspicuity) was defined as how well each sign being evaluated stood out as a whole. The brightness scale ranged from 1 (not bright) to 7 (very bright). Legibility was defined as how easy it was to read each sign. The scale ranged from 1 (not legible) to 7 (very legible). Adequacy was defined as how well each sign informed the observer and whether the sign could be used comfortably. The intent was to obtain observers' subjective evaluations of the acceptability of a sign without reference to other signs. The adequacy scale ranged from 1 (not adequate) to 7 (very adequate). Glare was defined as how shiny each sign was or how much reflection of unwanted light there was. The glare scale ranged from 1 (no glare) to 7 (excessive glare).

Finally, for each pair of signs the observers were asked to indicate which sign they preferred: left, right, or no preference; thus single signs did not receive preference ratings.

Analysis

For each observer, an average rating for each of the four 7-point scales was calculated for signs of the same combination of sheeting materials. Average ratings for the left and right signs were calculated separately. These data were submitted to within-subjects analyses of variance (ANOVAs). Signs in the left and right overhead positions were analyzed separately. In each analysis the material combinations were treated as the independent variables and the observers' responses on the four rating scales were treated as the dependent variables.

The assumption underlying the averaging of observers' ratings for signs of the same sheeting-material combination while controlling for overhead position (e.g., left, HI/HI) was that a measure would be obtained that was a more stable indicator of the performance of the sign combination across the driving environment of the test than would be obtained by analyzing each sign individually.

The data in Tables 2 and 3 give the mean ratings for the measures of perceived brightness, legibility, adequacy, and glare for the sheeting materials occurring, respectively, in the left and right overhead positions. The ANOVAs for the signs in the left overhead position indicated that observers did not perceive differences between the signs in terms of brightness, legibility, or adequacy ($p > 0.05$). However, a significant difference occurred for the measure of glare ($p < 0.001$). A Newman-Keuls test,

TABLE 2 Overhead Guide Signs in Left Position

	HI/HI	HI/EG	SEG/SEG	EG/EG
Brightness	5.27	4.98	5.00	4.89
Legibility	5.08	5.20	5.37	5.32
Adequacy	4.93	5.20	5.24	5.26
Glare[a]	4.57	3.67	3.66	3.37

Note: Data for 19 observers.
[a] $p < 0.001$.

TABLE 3 Overhead Guide Signs in Right Position

	HI/HI	HI/EG	SEG/EG	EG/EG
Brightness[a]	5.23	5.18	3.71	4.53
Legibility[a]	5.21	5.86	4.50	5.00
Adequacy[a]	5.05	5.55	4.24	4.74
Glare[a]	4.09	2.87	3.08	3.82

Note: Data for 19 observers.
[a] $p < 0.001$.

carried out to determine where differences between signs existed, indicated that the HI/HI combination was judged to have significantly more glare than any of the other three combinations ($p < 0.05$, for all combinations).

The ANOVAs for the signs in the right overhead position revealed significant differences in observers' judgments for all four measures ($p < 0.001$, for all measures). Newman-Keuls tests were performed to further explore these differences. The HI/HI and HI/EG signs were judged to be better than the SEG/EG and EG/EG signs ($p < 0.05$), whereas for both legibility and adequacy the HI/EG signs were judged to be superior to the other three sheeting material combinations ($p < 0.05$). As with the left overhead signs, the HI/HI (and EG/EG) signs were rated as having more glare than the HI/EG and SEG/EG signs ($p < 0.05$).

The foregoing analyses indicate that observers prefer signs with an HI foreground (legend and border) and an EG background, especially for signs in the right overhead position. Additional analyses were conducted to further explore this interpretation.

Because the average age of the HI/HI signs was higher than that of the other four combinations (see Objectives and Scope section), the data were reanalyzed excluding the data for HI/HI signs that had been erected more than 1 year before the study. The average age of the HI/HI signs in these analyses was 7.5 months for those in the left position ($n = 2$ signs) and 8.25 months for those in the right position ($n = 4$ signs). The results mirrored the results of the previously reported analyses. That is, in the left position the HI/HI signs were rated as having more glare than the other three signs; in the right position the HI/EG combination was perceived to be more legible and adequate than the other three signs, and again HI/HI was judged to have more glare. (All of the foregoing comparisons were statistically reliable with $p < 0.05$.) The mean ratings for the left and right overhead signs are given in Tables 4 and 5, respectively.

Data were collected in the test that required observers to indicate directly a preference between pairs of signs (not all of the possible combinations of pairs of signs were represented in the study). The preference judgments between pairs of signs of different sheeting materials were analyzed by using χ^2 tests. The preference ratings involving HI/EG signs consistently favored the HI/EG sign. Specifically, HI/EG signs were preferred over HI/HI signs when the HI/EG sign was in the left position (and the HI/HI

TABLE 4 Overhead Guide Signs Corrected for Age in Left Position

	HI/HI	HI/EG	SEG/SEG	EG/EG
Brightness	4.95	4.98	5.00	4.89
Legibility	5.16	5.20	5.37	5.32
Adequacy	5.21	5.20	5.24	5.26
Glare[a]	4.11	3.67	3.66	3.37

Note: Data are for 19 observers.
[a] $p < 0.05$.

TABLE 5 Overhead Guide Signs Corrected for Age in Right Position

	HI/HI	HI/EG	SEG/EG	EG/EG
Brightness[a]	5.32	5.18	3.71	4.53
Legibility[a]	5.18	5.86	4.50	5.00
Adequacy[a]	5.14	5.55	4.24	4.74
Glare[a]	3.97	2.87	3.08	3.82

Note: Data are for 19 observers.
[a] $p < 0.001$.

in the right) ($p < 0.001$, for three comparisons) as well as when the positions were reversed ($p < 0.001$, for one comparison). Also, in the single comparison of an HI/EG sign in the left position with an SEG/EG sign in the right, the HI/EG sign was preferred ($p < 0.001$). The HI/HI sign in the left position was preferred over both the SEG/EG sign ($p < 0.05$) and the EG/EG signs ($p < 0.05$) in the right position (both tests involved one comparison). However, no systematic preference was expressed between HI/HI signs in the right position and SEG/SEG signs in the left position (involved two comparisons).

The several analyses consistently indicate that the observers had a marked preference for HI/EG signs. This conclusion is especially warranted when HI/EG signs are compared with HI/HI signs, because both types of sheeting material combinations were adequately represented in the study.

LUMINANCE MEASUREMENTS

All measurements were made with a Pritchard Spectrum Photometer Model No. 1980A, which was able to measure target areas contained within an angle range (angle of acceptance) from 0° 2' to 3° 00'. The instrument panel has a digital readout in candelas per square meter (cd/m²), with the sensitive range (measuring span) from 10^{-4} to 10^{8} cd/m² with photopic color correction, calibrated within ±4 percent of reading or 2 percent full-scale accuracy, whichever is greater. The smallest 2-minute angle of acceptance was selected because the target area contained within this angle could fit onto the narrow width of the letters of the legend at the maximum distance of 50 m (164 ft). The instrument was mounted with a specially designed mount on the passenger side with lens height at the eye level. Two operators carried out the measurements. One (passenger) aligned the optical head with the object in the field of view, while the other (driver) recorded the measurements.

The test vehicle was a standard domestic Chrysler Panel Van 198. Before the readings the windshield and headlamp surfaces were cleaned. The vehicle was positioned on the right shoulder of the roadway 50 m in front of the sign.

The background luminances were measured at four corners within the borders of available space and in the center of the sign. The sign-legend luminance was taken on the crown and arrow (if present) and the first and the last letter of each string of the message.

The luminance readings were taken while the sign was illuminated with the combination of ambient illumination (roadway lighting and so forth), vehicles moving on the roadway, and the headlights of the test vehicle (low beams only).

The average luminance values (cd/m²) for each sample reading were grouped by facing material type for background and legend of each sign for each type of facing material. Measurements could not be undertaken safely for all the signs used in the subjective evaluation because of the roadway geometrics of some sign locations.

Traditionally, contrast and luminance levels are among the parameters considered as major factors affecting legibility of a sign. These parameters, in turn, are dependent on reflectivity characteristics, color, and size of legend.

The definition of contrast used in this paper is based on the following requirements and rationale, where contrast equals the luminance ratio:

$$C = L_2/L_1 \qquad (1)$$

where

C = contrast,
L_1 = background luminance, and
L_2 = legend luminance.

The relationship between legibility distance and the legend-to-background-luminance ratios has been developed by Forbes et al. (1,2). This relationship defines, for white legend and green background, that with 20/20 vision and position between the light source and the sign, a maximum legibility distance of 6.0 to 7.2 m/cm (50 to 60 ft/in.) of letter height can be generally obtained. To achieve such conditions, typical legend-to-background-luminance ratios lie within the range of 6:1 to 13:1.

However, Forbes et al. (1,2) also state that in practice the luminance ratios cannot be achieved and the ratios that can be expected will be within the range of 3:1 to 7:1, which when translated into legibility distance correspond to 5.4 to 6.0 m/cm (45 to 50 ft/in.) of letter height.

The average contrast ratios for the HI/HI and HI/EG material combinations are given in the following table:

	Position	
Material	Left	Right
HI/HI	5.5	4.9
HI/EG	5.8	6.4

The test results indicate that either HI/HI or HI/EG will provide satisfactory contrast ratios for legibility, with high-intensity foreground (legend and border) on engineering-grade background giving somewhat better contrast ratios. This finding supports the subjective test results. The conclusion drawn from the luminance measurements is that the substitution of engineering-grade background for high-intensity background will not reduce sign legibility, provided that the legend of the sign is made of high-intensity reflective material.

COST ANALYSIS

For the cost assessment based on empirical experience over the past several years, the following assumptions were made:

1. Cost of HI/HI sign, including material and labor, equals $70.00/m² ($6.50/ft²);

2. Cost of HI/EG sign, including material and labor, equals $43.00/m² ($4.00/ft²);
3. Area of freeway overhead guide signs in Ontario on provincial freeways equals 13,940 m² (150,000 ft²);
4. Expected life of engineering-grade material equals 10 years; and
5. Expected life of high-intensity material (when used as background) equals 5 years.

The result of using high-intensity material for both foreground and background and having to refurbish signs every 5 years (on average) would be an annual refurbishing cost of about $195,000 per year. When using high-intensity foreground and engineering-grade background, there are two extreme cases to be considered:

1. The life of high-intensity material when used as foreground is 5 years, which necessitates refurbishing of the complete sign at the 5-year point; the annual refurbishing cost is $120,000 per year.
2. The life of high-intensity material when used as foreground is 10 years, which necessitates complete sign refurbishing after 10 years; the annual refurbishing cost is $60,000 per year.

Annual savings resulting from use of HI/EG rather than HI/HI on overhead freeway guide signs would appear to lie between $75,000 and $135,000. Other possibilities, falling between these two extremes, include high-intensity foreground life between 5 and 10 years; engineering-grade background life less than 10 years; and partial refurbishing before 10 years without having to scrap useful life of the engineering-grade background.

SUMMARY AND CONCLUSIONS

Three sources of information were employed in the comparative evaluation of various combinations of reflective sheeting on freeway overhead guide signs: observers' judgments, luminance measures, and cost analysis. These three sources converged in recommending, on balance, the use of high-intensity foreground (legends and borders) on engineering-grade background for freeway overhead guide signs.

Observers favored the HI/EG combination both in rating the features of these signs (more legible, more adequate, and less glare) and in consistently choosing the HI/EG combination over each of the other combinations when stating their preference judgments. The analysis of cost between HI/HI and HI/EG clearly favors the latter combination, and luminance measurements indicate that HI/EG provides contrast ratios for legibility that are at least as satisfactory as those for HI/HI.

The present study has some limitations that should be addressed. First, as mentioned previously, the evaluation focused on a comparison between high-intensity and engineering-grade backgrounds, when both had high-intensity foregrounds. The combination of HI/SEG was not represented, although it could be considered as having the potential to be a satisfactory overhead guide sign.

A second limitation results from the practical considerations of testing under clear weather conditions. Previous research (3) has identified the importance of testing under degraded visual conditions when assessing a sign's content (i.e., verbal versus symbolic messages). It would appear reasonable to extend this concern to evaluations of sign sheeting reflectivity. It has been suggested, however, that in wet or rainy conditions, sign conspicuity may be improved because of the additional light reflected from the wet roadway onto the sign. Currently there is no evidence to suggest that a particular combination of sheeting materials would perform better under degraded visual conditions.

Finally, although the sample of observers did range in age from 20 to 60, there were not enough observers in the different age categories (e.g., younger than 30, 31 to 50) to analyze for the effect of age. Research by Sivak et al. (4,5) has identified age as an important variable in tests of nighttime legibility of signs. Specifically, this research indicates that younger observers enjoy an advantage over older observers in distance of legibility. However, the more recent research by these investigators (5) found that the age-related decrement in performance was eliminated with increased contrast ratios for the letter-background combinations of signs. This latter result appears consistent with favoring an HI/EG combination with its higher contrast ratio over an HI/HI combination.

In conclusion, any decision on use of sign materials should be based on an evaluation of observers' reactions to the sign and its cost-effectiveness. It is recognized that certain potential limitations exist on the evaluations of the observers, and it is acknowledged that cost-effectiveness is based on a constantly changing equation that is affected by initial cost, product durability and life, and reflectivity. Having addressed both issues, this study reached a recommendation about sign sheeting. Specifically, pending further research, it is recommended that for freeway overhead guide signs a high-intensity foreground (legend and border) on an engineering-grade background be used.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the assistance and support of the staff in the Traffic Surveys Section for serving as drivers, Chris Blamey for helping design the route and train the drivers, Susan Williams for analyzing the data and her constructive contributions in the preparation of this paper, and, especially, to the employees of MTC who volunteered their time to serve as observers in the study.

REFERENCES

1. T.W. Forbes, B.B. Saari, W.H. Greenwood, J.G. Goldblatt, and T.E. Hill. Luminance and Contrast Requirements for Legibility and Visibility of Highway Signs. In Transportation Research Record 562, TRB, National Research Council, Washington, D.C., 1976, pp. 59-72.
2. T.W. Forbes. Luminance and Contrasts for Sign Legibility and Color Recognition. In Transportation Research Record 611, TRB, National Research Council, Washington, D.C., 1976, pp. 17-25.
3. J.G. Ellis and R.E. Dewar. Rapid Comprehension of Verbal and Symbolic Traffic Sign Messages. Human Factors, Vol. 21, No. 2, 1979, pp. 161-168.
4. M. Sivak, P.L. Olson, and L.A. Pastalan. Effect of Driver's Age on Nighttime Legibility of Highway Signs. Human Factors, Vol. 23, 1981, pp. 59-64.
5. M. Sivak and P.L. Olson. Nighttime Legibility of Traffic Signs: Conditions Eliminating the Effects of Driver Age and Disability Glare. Accident Analysis and Prevention, Vol. 14, No. 2, 1982, pp. 87-93.

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

Effectiveness of Wildlife Warning Reflectors in Reducing Deer-Vehicle Accidents in Washington State

JAMES A. SCHAFER, STEPHEN PENLAND, and WILLIAM P. CARR

ABSTRACT

The effectiveness of Swareflex Wildlife Reflectors in reducing deer-vehicle collision rates was tested on SR-395 in eastern Washington State, on which high mortality rates of white-tailed deer (Odocoileus virginianus) had previously been recorded. Reflectors were placed in four test sections and alternately covered and uncovered at regular intervals during the late fall to early spring period from 1981 to 1984. During this period 52 deer were killed at night in test sections when the reflectors were covered and 6 deer were killed at night when the reflectors were uncovered. This difference in deer-vehicle collision rates between the covered and uncovered periods is significant ($p < 0.005$), which indicates that the reflectors were effective on this highway during this time period.

Collisions between deer and automobiles produce a substantial economic cost through damage to vehicles, the loss of a valuable wildlife resource, and human injuries or fatalities. Since 1977, 3,142 deer-vehicle accidents have been recorded by the Washington State Department of Transportation (WSDOT). High accident rates have been estimated in other states, including 3,000 in Iowa in 1978 (1), 4,900 in Colorado in 1968 (2), an average of more than 12,600 annually between 1972 and 1976 in Michigan (3), and 22,000 annually in the early 1970s in Pennsylvania (4).

Pils and Martin (5) and Reed et al. (6) estimated that the average cost of vehicular damage in these kinds of collisions was $500 in 1978. Washington State Patrol records indicate that 108 reported deer-vehicle collisions resulted in $82,000 in automobile damages and six human injuries on just one 30-mile stretch of SR-395 in eastern Washington since 1977. Adding the costs of human deaths and injuries, Hanson (7) estimated that each deer-vehicle accident cost $730. The economic value of each deer killed is more difficult to quantify (8). Reed et al. (6), using a damage award from a Colorado District Court, placed the economic loss of a deer at $350 in 1976. Hartman (9) and Norman (10) placed a deer's value at more than twice this figure based on hunting expenditures alone. Clearly, the 200,000 annual deer-vehicle collisions on America's highways (11) result in the loss of many millions of dollars.

A new reflector system designed to reduce the number of deer-vehicle accidents has been developed in Austria. This system, called Swareflex Wildlife Reflectors (Figure 1), consists of a series of 6.5 x 2-in. red reflectors mounted along the roadway (Figure 2). Light from the headlights of an approaching automobile is reflected at right angles to the roadway by the reflectors, creating an "optical

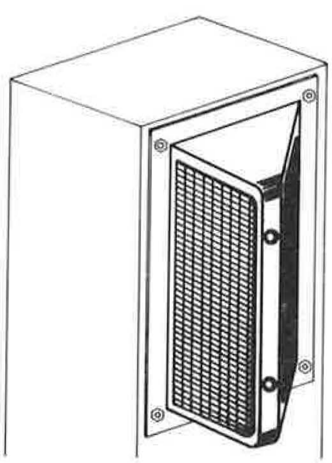


FIGURE 1 Wildlife warning reflector.

FIGURE 2 Reflector installation.

fence" that presumably causes deer to remain motionless until the automobile has passed and the optical fence has collapsed. Unfortunately, most tests of the effectiveness of the Swareflex Reflectors have consisted of before-and-after comparisons of deer kills that are confounded by variations of annual weather patterns, deer population densities, and traffic patterns. WSDOT used an experimental covered-uncovered design developed with the help of Charles T. Robbins of Washington State University that allows a valid statistical evaluation of the Swareflex Reflector system.

METHODS

Four test sections were established along SR-395 in an arid transitional ponderosa pine forest-grassland zone north of Spokane, Washington (Table 1). Each test section was placed in an area with high deer-vehicle accident rates. The sections ranged from 0.45 to 0.68 mile in length. Reflectors were placed at 66-ft intervals along straight road sections and 33-ft intervals on curves on both sides of the roadway, as suggested by the manufacturer.

TABLE 1 Locations and Number of Deer Killed in Test Sections

Section	Milepost	Total Miles	No. of Deer Killed	
			Covered	Uncovered
A	214.40–214.90	0.50	11	0
B	217.26–217.94	0.68	8	1
C	218.53–218.98	0.45	17	2
D	219.85–219.97	0.67	16	3
	220.05–220.13			
	220.26–220.44			
	220.52–220.60			
	220.62–220.76			
	220.93–221.05			

The reflectors in each test section were alternately covered and uncovered at 1-week intervals between mid-October and mid-April each year from February 1981 to April 1984. The covered-uncovered period was extended to 2-week intervals after December 1982. Alternate test sections were paired so that reflectors in each pair were covered while reflectors in their adjacent sections were uncovered, and vice versa.

The highway was traveled daily by WSDOT maintenance personnel. The milepost location, estimated time of kill, and the covered-uncovered status of the Swareflex Reflectors were recorded for each dead deer found along the highway. A paired t-test (12) was used to compare the number of deer killed at night during periods when reflectors were covered with the number killed at night during periods when reflectors were uncovered.

RESULTS

A total of 1,619 deer were killed on state highways from 1981 through May 1984. This total included 594 (37 percent) that were killed on SR-395. Seventy percent of the 801 deer killed statewide at known times of the day were killed during the nighttime hours.

The number of deer killed on SR-395 during the mid-October to mid-April test period since 1981 was 363, or 61 percent of the total number killed on that highway. Seventy-three (20 percent) were killed within the 2.3 miles of the test sections. The 138 deer killed outside the test sections at known times of the day included 114 (83 percent) that were killed at night and 24 (17 percent) that were killed during the day.

Fifty-eight deer were killed at night in the test sections during the test period (Table 1). These included 56 white-tailed deer (*Odocoileus virginianus*) and 2 mule (*O. hemionus*) deer. Fifty-two deer (90 percent) were killed when the reflectors were covered, and six (10 percent) were killed when the reflectors were uncovered. The difference between the number of deer killed when the reflectors were covered and the number killed when the reflectors were uncovered is statistically significant ($p < 0.005$).

DISCUSSION OF RESULTS

Swareflex Reflectors have usually been evaluated by comparing the number of deer killed along roadways after reflector installation with deer kills recorded before reflector installation. These comparisons have usually revealed a reduction in deer-vehicle collisions after reflector installation (citations from personal communication with Strieter Corporation). But annual variations of considerable magnitude exist in rates of deer-vehicle collisions (Table 2), probably because of changing deer population densities, changing traffic patterns, differences in weather that affect deer movement, or other factors (3,13-16). These variations obscure the relationship between reflectors and deer-vehicle collision rates when comparisons are made over periods of time. The use of an alternating present-absent study design eliminates the effects of these large-scale variations and allows a statistical evaluation of reflector effectiveness. A present-absent study design was used by Woodard et al. (17) for 24 weeks in Colorado. Because 11 deer were killed on a 1-mile test section when the reflectors were present compared with 8 deer killed when the reflectors were absent, they concluded that the Swareflex Reflectors were not effective. However, they did not describe the method of censusing dead deer nor did they specify whether the deer were killed only at night.

Polished stainless-steel mirrors, often called Van de Ree reflectors, have also been tested for their ability to reduce deer-vehicle collisions. Gilbert (18) attempted to reduce the variations inherent in time comparisons by using Van de Ree mirrors in twelve 0.5-mile randomly located sections along a 14.8-mile freeway in Maine. After 3 years, four deer had been killed in mirrored sections and three had been killed in nonmirrored sections. This small sample size did not permit a statistical test

TABLE 2 Annual Numbers of Deer-Vehicle Collisions in Washington State

Location	No. of Deer-Vehicle Collisions							
	1977	1978	1979	1980	1981	1982	1983	Total
SR-395	99	240	174	119	185	187	308	1,312
Eastern Washington	168	394	246	227	218	220	361	1,834
Western Washington	263	241	201	124	87	184	208	1,308
State total	431	635	447	351	305	404	569	3,142

of mirror effectiveness. Even if the sample size had been sufficient, the interpretation of data from randomly located test sections would still be plagued by the nonrandom distribution of deer because of differences in topography and resource availability. Other tests on Van de Ree mirrors have generally employed a before-and-after study design (18). Most have concluded that the mirrors were ineffective, although one test in Maine offered a qualified success and one in the Netherlands reported a 100 percent reduction in the number of deer killed during a 4-year period.

Although WSDOT's test of the Swareflex Reflectors was conducted during the late fall, winter, and early spring months, the distribution of deer-vehicle collisions reveals only a modest increase in the number of deer killed in February and March on SR-395 and other highways in eastern Washington (Figure 3). The number of deer killed by cars in western Washington peaks in the summer months and is lowest during the winter. Reports from other states have indicated that the most deer activity along highways and the highest mortality on highways occur in late fall and, to a lesser extent, in spring (3,4,14,19,20). Reilly and Green (15) found a pronounced late winter-early spring peak in highway mortality of deer in northern Michigan that was in contrast to the fall peak of highway mortality in other parts of the state. Case (16) reported a peak in highway mortality of deer in Nebraska during May and June, and a somewhat smaller increase during October and November.

The manufacturer of the Swareflex Reflectors claims that the red color of the reflectors initiates an instinctive "freezing" response in deer. Evidence for this functional response to red color has been given by Backhaus (21) and discussed by Koenig (22) and Weis (23), although Severinghaus and Cheatum (24) stated that deer are color-blind. Whether the red color or simply the point source of light produces the functional response, the reflectors are effective only during the hours of darkness. Of the deer killed by vehicles at known times in Washington State, most are killed at night. A similar majority of the deer-vehicle accidents in other states also occur after sunset (2,3,20).

SUMMARY AND CONCLUSIONS

The economic cost of deer-vehicle collisions warrants consideration of effective preventive measures. The results and interpretations of previous studies of the effectiveness of deer mirrors have been hampered by small sample sizes and by influences of large-scale environmental factors on deer-vehicle collision rates over time when before-and-after comparisons are made. WSDOT employed an alternating cover-uncover study design to test the effectiveness of Swareflex Reflectors in an area

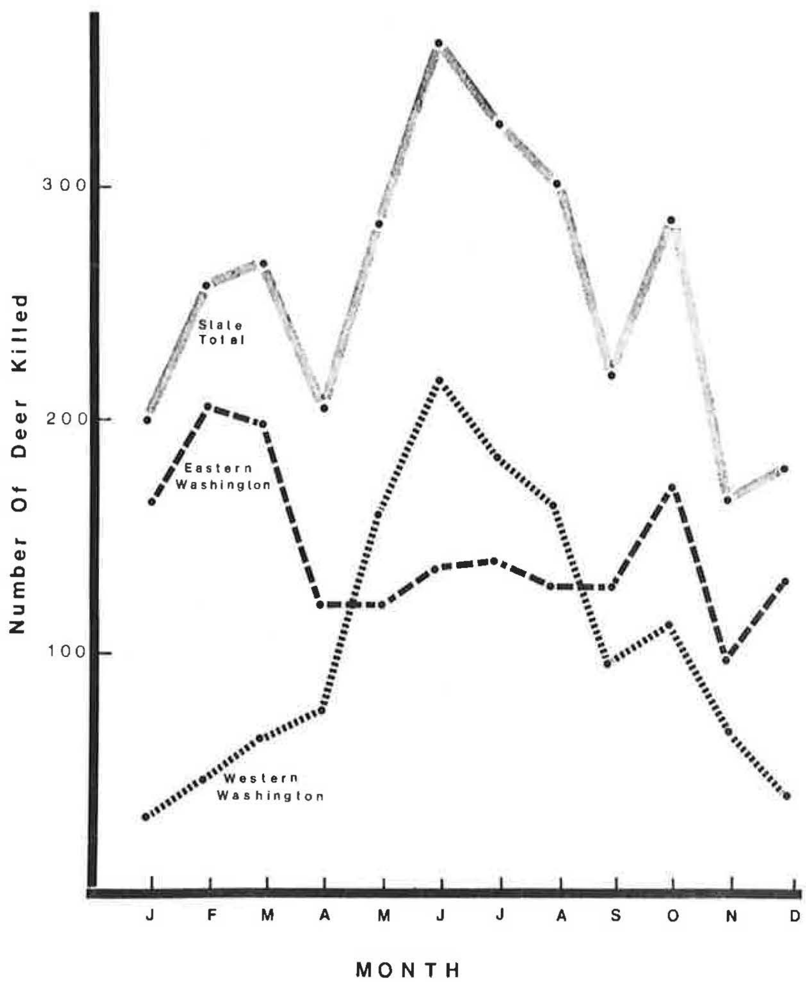


FIGURE 3 Average monthly distribution of deer-vehicle collisions between 1977 and 1983 in Washington State.

with historically high rates of deer-vehicle collisions. After 3 years, the reduction in the number of deer killed when the reflectors were uncovered was statistically significant. The Swareflex Reflectors were effective in reducing deer-vehicle collisions on this state highway in Washington.

ACKNOWLEDGMENTS

The authors gratefully acknowledge D.G. Pierce and the highway maintenance personnel of the Coleville maintenance office for recording deer kills and changing reflector covers during the study period; their consistent cooperation made the study possible. Charles T. Robbins provided suggestions for the study design and locations of test sections. Computer services were provided by O.G. Wells and L.K. Egge. Graphic services were provided by Z.M. Olson.

REFERENCES

1. L. Gladfelter. Stop Signs for Deer. Iowa Conservationist, Vol. 39, No. 8, 1980, pp. 10-11.
2. G.T. Meyers. Deer-Auto Accidents--Serious Business. Colorado Outdoors, Vol. 18, No. 3, 1969, pp. 38-40.
3. D.A. Arnold. Characteristic and Cost of Highway Deer Kills. In Proc., John S. Wright Forest Conference (C. Kirkpatrick, ed.), Department of Forestry and Natural Resources and Indiana Cooperative Extension Service, Purdue University, West Lafayette, Ind., 1978, pp. 92-101.
4. M.J. Puglisi, J.S. Lindzey, and E.D. Bellis. Factors Associated with Highway Mortality of White-Tailed Deer. Journal of Wildlife Management, Vol. 38, No. 3, 1974, pp. 799-807.
5. C.M. Pils and M.A. Martin. The Cost and Chronollogy of Wisconsin Deer-Vehicle Collisions. Research Report 103. Wisconsin Department of Natural Resources, Madison, 1979, 5 pp.
6. D.F. Reed, T.D. Bech, and T.N. Woodard. Methods of Reducing Deer-Vehicle Accidents: Benefit-Cost Analysis. Wildlife Society Bulletin, Vol. 10, 1982, pp. 349-354.
7. C.D. Hanson. Social Costs of Michigan's Deer Habitat Improvement Program. Ph.D. dissertation. University of Michigan, Ann Arbor, 1977, 195 pp.
8. W.A. Langford and D.J. Cocheba. The Wildlife Valuation Problem: A Critical Review of Economic Approaches. Occasional Paper 37. Canadian Wildlife Service, 1978, 37 pp.
9. F.E. Hartman. Hunting is Big Business. Pennsylvania Game News, Vol. 44, 1973, pp. 26-31.
10. R. Norman. Using Wildlife Values in Benefit-Cost Analysis and Mitigation of Wildlife Losses. Proc., International Association of Game, Fish, and Conservation Commissioners, Vol. 56, 1975, pp. 119-128.
11. L. Williamson. Reflectors Reduce Deer-Auto Collisions. Outdoor News Bulletin, Vol. 34, No. 2, 1980, p. 2.
12. R. Sokal and F. Rohlf. Biometry. W.H. Freeman and Company, San Francisco, 1969, 776 pp.
13. L.M. Queal. Effectiveness of Roadside Mirrors in Controlling Deer-Car Accidents. Report 137. Michigan Department of Conservation Resources and Development, Lansing, 1968.
14. E.D. Bellis and H.B. Graves. Deer Mortality on a Pennsylvania Interstate Highway. Journal of Wildlife Management, Vol. 35, No. 2, 1971, pp. 232-237.
15. R.E. Reilly and H.E. Green. Deer Mortality on a Michigan Interstate Highway. Journal of Wildlife Management, Vol. 38, No. 1, 1974, pp. 16-19.
16. R.M. Case. Interstate Highway Road-Killed Animals: A Data Source for Biologists. Wildlife Society Bulletin, Vol. 6, No. 1, 1978, pp. 8-13.
17. T.N. Woodard, D.F. Reed, and T.M. Pojar. Effectiveness of Swareflex Wildlife Warning Reflectors in Reducing Deer-Vehicle Accidents. Internal Report. Colorado Division of Wildlife, Denver, 1973, 5 pp.
18. J.R. Gilbert. Evaluation of Deer Mirrors for Reducing Deer-Vehicle Collisions. Report FHWA/RD-82/061. FHWA, U.S. Department of Transportation, 1982.
19. B. Carbaugh, J.P. Vaughan, E.D. Bellis, and H.B. Graves. Distribution and Activity of White-Tailed Deer Along an Interstate Highway. Journal of Wildlife Management, Vol. 39, No. 3, 1975, pp. 570-581.
20. R.E. Allen and D.R. McCullough. Deer-Car Accidents in Southern Michigan. Journal of Wildlife Management, Vol. 40, No. 2, 1976, pp. 317-325.
21. D. Backhaus. Experimental Investigations on the Activity of Vision and Color Vision in Some Hoofed Animals. Zeitschrift fuer Tierpsychologie, Vol. 16, No. 4, 1969, pp. 465-467 (in German, with English summary).
22. O. Koenig. Eye-Point of Vision. Strieter Corporation, Austria, 1974.
23. J. Weis. Deer Do See Color! Outdoor Life, Vol. 167, No. 3, 1981, p. 64.
24. C.W. Severinghaus and E.L. Cheatum. Life and Times of the White-Tailed Deer. In The Deer of North America (W.P. Taylor, ed.), The Stackpole Company, Harrisburg, Pa., 1956, pp. 57-186.

This paper reflects the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Washington State Department of Transportation; no endorsement of any product is implied or intended. This paper does not constitute a standard, specification, or regulation.

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

Abridgment

Monitoring and Evaluation of High-Type Railroad Crossing Surfaces

DEAN A. MAURER

ABSTRACT

High-type railroad crossing surface improvements have resulted in the replacement of traditional asphalt and timber crossing surfaces with specially designed durable materials and replacement or improvement of the track structure and the supporting subgrade. In this paper the various types of proprietary surfacing products installed in Pennsylvania through 1983 are summarized and the construction and design problems associated with each product are discussed. To remedy construction-related problems, which are considered to be significant, detailed guidelines for site preparation and installation of high-type railroad crossing surfaces are proposed.

A common problem that has occurred for many years on highway systems universally has been the existence of rough and dangerous railroad crossings (Figure 1). Even though the predominant causes for the problem--excessive loading of the subgrade and inadequate drainage--have been recognized for a long time, only a limited effort has been made to correct the situation until the past decade. Starting in 1973, Section 203 of the Federal Highway Safety Act provided special funding for the construction of high-type surfaces at selected railroad-highway grade crossings. This was the start of a comprehensive program to improve grade crossings across the United States. Other improvements funded by the program include provisions for upgrading protective or warning devices and the elimination of crossings where abandonment of rail service has occurred.

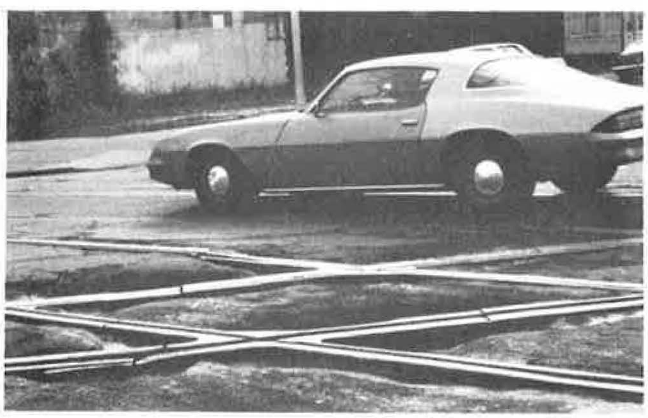


FIGURE 1 A high-type crossing surface is desirable for a rough and dangerous crossing such as this.

High-type surface improvements have resulted in the replacement of traditional asphalt and timber crossing surfaces with specially designed durable materials and replacement or improvement of the track structure and the supporting subgrade. An essential aspect of subgrade improvement has been the provision for adequate drainage. Because most of the developed surfacing products have been of a proprietary nature, their use on federally funded projects requires, by law, evaluation of performance. The Pennsylvania Department of Transportation (PennDOT), participating in the program established by the Federal Highway Safety Act, developed Research Project 77-21 in 1977. This project, entitled "High-Type Railroad Crossing Surface Monitoring and Evaluation," is PennDOT's commitment to evaluating and reporting the performance of proprietary materials used for constructing improved grade crossings.

The interim report presented herein is primarily an update of PennDOT's experience between 1978 and 1983 (using Section 203 funds). The focus of this report is on summarizing the various types of proprietary surfacing products installed through 1983 and construction and design-related problems identified with such installations. Guidelines for site preparation and installation of high-type railroad crossing surfaces are recommended for inclusion in contracts for new crossings to improve the performance of installation. An outline of Pennsylvania's new guidelines is provided for consideration.

I. Types of high-type surfaces installed
   A. Partial depth panels with timber shims
      1. "Super cushion" rubber pads (Goodyear Tire and Rubber Company): steel plate reinforced-rubber (elastomeric) pads secured to ties by lug bolts (drive spike)
      2. "Parkco" rubber pads (Park Rubber Company): steel plate reinforced-rubber pads secured in place by steel tensioned cables
   B. Full-depth sections (shimless)
      1. "Gen-Trac" elastomeric grade crossing (General Tire and Rubber Company): steel arch reinforced-rubber units secured to ties by lug bolts (drive spikes)
      2. "Cobra-X" grade crossing modules (Railroad Friction Products Corporation): high-density polyethylene ejection-molded modules secured to ties by lug bolts (drive spikes)
      3. "True Temper" grade-crossing modules (True Temper Corporation): high-density polyethylene structural foam, pressure-molded modules secured to ties by lug bolts (drive spikes)

4. "Omni" shimless grade crossing (precured RDF Tirefill, Inc., now Omni Rubber Products): rubber panels manufactured from 100 percent rubber material from tire retreading process secured to ties by self-tapping timber bolts

C. Relative use in the United States since 1978 (approximate), by manufacturer and number of crossing sites installed

1. Goodyear, 23
2. Parkco, 19
3. Gen-Trac, 8
4. Cobra-X, 3
5. True Temper, 1
6. Omni, 1

II. General problems associated with all types of crossing installations observed (* = widespread occurrence or a major problem)

A. *Failed replacement approach pavement and joint with crossing surface (cracks and settlement) due to

1. Inadequate compaction of subgrade, ballast, and pavement material in the area from the end of tie to the existing pavement
2. Failure to install header board
3. Misaligned and damaged header boards
4. Failure to seal or maintain pavement-crossing joint with rubberized asphalt sealant (see Figure 2)

FIGURE 2 Failure to seal joint resulted in pavement cracking. Note that one-half is sealed but has no header board and is performing better.

5. Inadequate existing pavement removed for crossing installation, creating a space too narrow to properly compact replacement materials (see Figures 3 and 4)
6. Improper crosstie length, creating voids and misalignment between header boards

B. *Crossing settlement causing poor transition and premature loss of riding comfort due to

1. Inadequate ballast depth or compaction or both under rails (see Figure 5)
2. Unstable subgrade (inadequate preliminary investigation by soils engineer)
3. Inadequate or improperly installed drainage system (see Figure 6)
4. Improper establishment of highway crossing elevation

C. Poor drainage of crossing area due to

1. Improper size of coarse aggregate for pipe backfill
2. Damaged pipe used and improperly installed

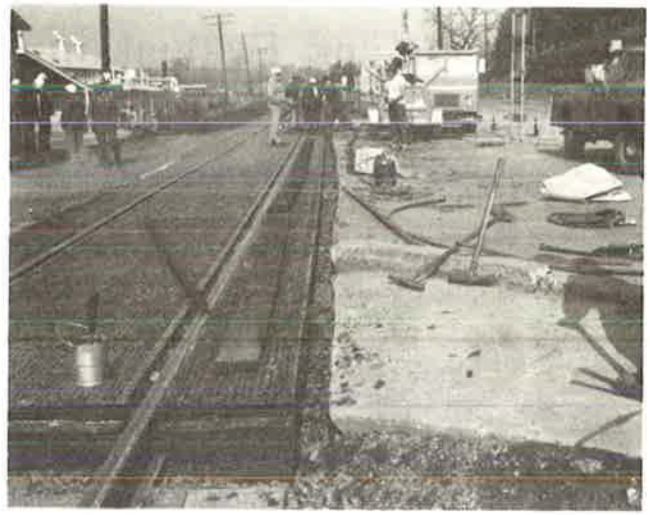


FIGURE 3 When insufficient pavement is removed for construction, it is improbable that the replacement material will be thoroughly compacted.

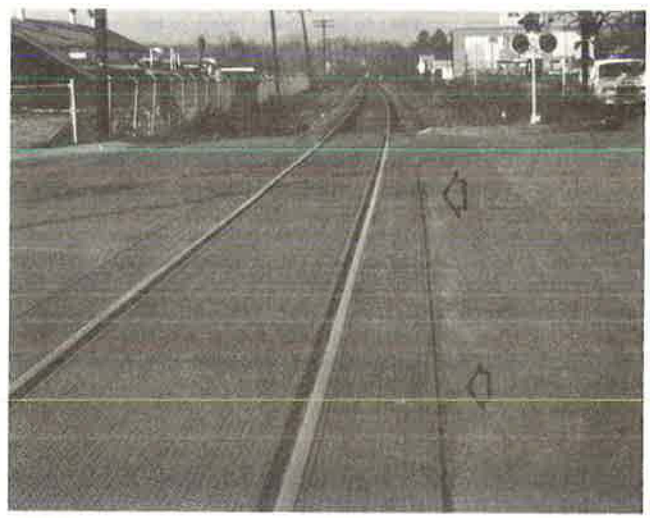


FIGURE 4 Condition of joint shown in Figure 3 after 3 years.

FIGURE 5 This method of compaction will not prevent settlement of track structure.

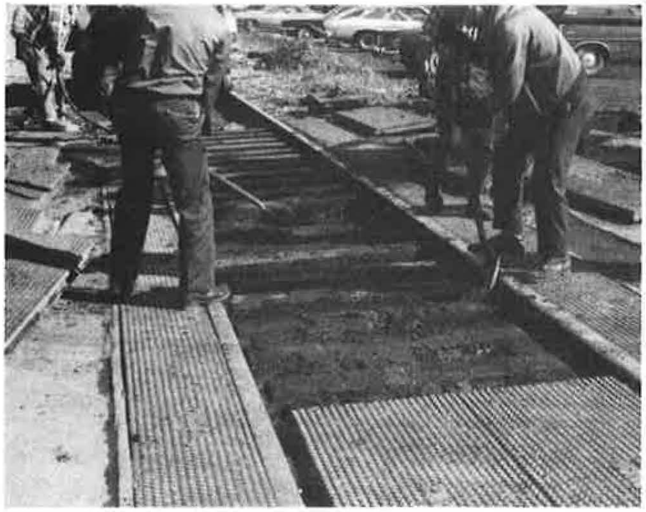


FIGURE 6 Because of a high water table and poor drainage design, crossing was completely clogged with saturated fines after only 3 years.

3. Pipe improperly sloped

4. *Heat-bonded geotextiles often chosen to wrap trenches, which trap water and do not provide planar flow

5. *Excess flow into crossing area when excessive debris and road dirt accumulate along the flange way; need for routine cleaning to maintain seal of high-type surface (see Figure 7)

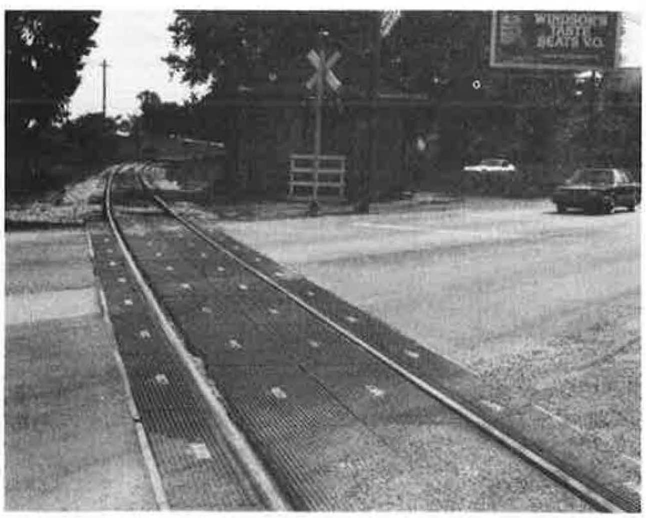


FIGURE 7 Without routine cleaning, excess highway anti-skid material has accumulated in shoulder area and in flange way.

D. *Inadequate geotextiles used for trackbed stabilization; lightweight and heat-bonded fabrics often selected due to inadequate specification in guidelines because minimal performance data available for geotextiles during initial guideline write-up

E. Damage to high-type crossing surface material due to

1. Dragging railroad equipment where end drag protection plates are missing

2. Digging by blades of highway snowplows

3. Same as C.4.--Deforms shape of crossing material

III. Specific problems associated with type of crossing (design)

A. Major failure of the high-type surface material caused by structural failure of the member or deterioration of material from abrasion (Cobra X, Omni, Gen-Trac)

B. *Loose or broken supporting wood shims (Parkco, Goodyear)

C. Failure of the mechanism securing the surface material to the track structure due to either overstressing or corrosion caused by poor drainage of track structure (Parkco tension cables)

D. Loss of numerous panel spike plugs or rubber caps (Goodyear, Gen-Trac)

E. Minor cracking in surface material (Gen-Trac)

DETAILED CONSTRUCTION GUIDELINES INDICATE IMPROVEMENT

Several projects in Pennsylvania have been constructed by using revised, detailed specifications for crossing installations designed to regulate construction procedures more closely. The specifications or guidelines place particular emphasis on compaction, drainage installation, and pavement-crossing joint construction (see Figures 8-12). Early evaluation of these crossings indicates longer satisfactory performance is expected relative to prior installations constructed with less stringent controls. An outline of these guidelines follows (an unabridged printing of the guidelines is available from the Information Center of the Pennsylvania Department of Transportation by requesting a copy of the report for Research Project 77-21):

I. General design requirements

A. Outline responsibility and involvement of all parties in contract

B. Highlight design criteria

II. Preconstruction coordination

A. Establish method of submission and acceptance of design proposal

FIGURE 8 A properly designed and installed header board serves several functions. It provides a uniform rigid wall against which to compact the approach.

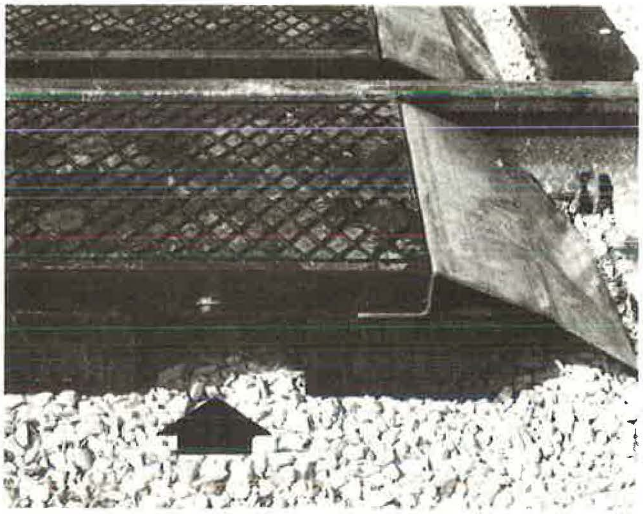


FIGURE 9 A header board placed full depth prevents loss of material in the tie cribs, which would eventually result in joint settlement.

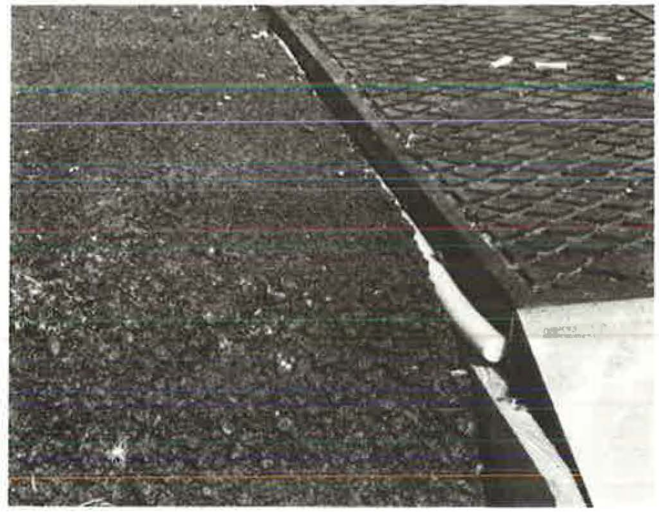


FIGURE 11 Backer rod is recommended because it forms desired shape and prevents adhesion on bottom of joint for proper movement.

B. Establish method of traffic control to be maintained during construction (a detour is always preferred)

C. Detail limits of inspection and criteria for acceptance of work

D. Attendance of all parties at an on-site preconstruction meeting usually promotes better cooperation between parties, provides better understanding of work to be accomplished, and eliminates surprise conditions

III. Construction requirements: provide detailed specifications for performance of each phase of work [* = providing an accompanying detailed sketch or specification is beneficial for these items (see Figure 13)]

A. Maintenance and protection of traffic

B. *Site preparation

C. *Drainage installation and subgrade stabilization

D. *Ballast replacement

E. *Track reconstruction

F. *Crossing surface installation

G. *Approach pavement-crossing joint construction

H. *Highway approach pavement reconstruction

IV. Maintenance requirements and warranties: provide details of post-construction requirements (if any) for each party (highway department, railroad, surface manufacturer, and so forth)

INTERIM CONCLUSIONS

1. Lack of quality control during installation and inappropriate or inadequate construction procedures are often the primary causes of premature failure of high-type grade crossings.

2. Detailed installation guidelines or specifications can effectively reduce premature failure of many grade crossings by providing a uniform measure of quality control during construction. Particular emphasis on compaction, drainage installation, and pavement-crossing joint construction is essential for long-term performance.

FIGURE 10 The top of a header board can also be designed to facilitate the construction of a flexible sealed joint between the pavement and the crossing.

FIGURE 12 Hot-poured rubberized sealant has performed best.

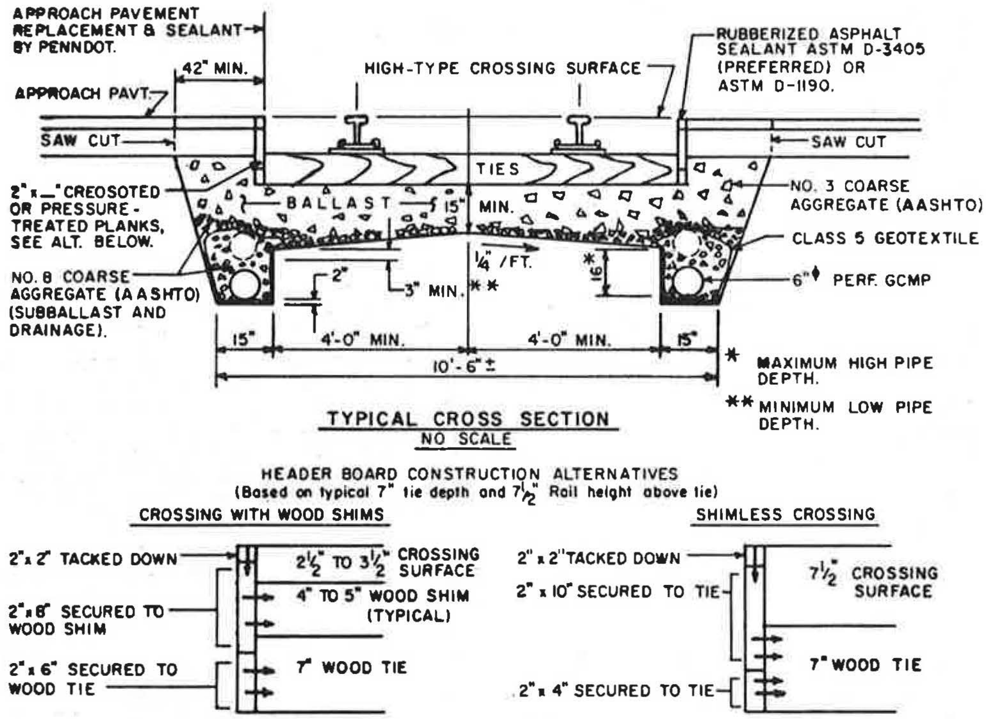


FIGURE 13 Typical detailed sketch.

3. Although most high-type crossing failures that have been evaluated indicate failure related to construction details, some failures indicate that the cause was compounded or even independently due to design factors related to structural support, type of connections, and material type. Continued evaluation is necessary to achieve complete documentation and to determine the relevance and frequency of the observed deficiencies.

4. Routine inspection to determine maintenance needs, such as sealing of pavement-crossing joints and cleaning debris from rail flange ways, is essential for the long-term performance of in-service crossings.

ACKNOWLEDGMENT

This paper is an abridgment of an interim report for Research Project 77-21, funded by FHWA and conducted by PennDOT.

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

Further Investigation of the Effectiveness of Warning Devices at Rail-Highway Grade Crossings

RONALD W. ECK and JOHN A. HALKIAS

ABSTRACT

The main objective of the study was to analyze the national inventory of the U.S. Department of Transportation-Association of American Railroads and the accident files of the Federal Railroad Administration to develop measures of effectiveness for the following rail-highway grade-crossing upgrade stratifications: (a) passive systems to flashing lights on single track, (b) passive systems to gates on single and multiple track, and (c) flashing lights to gates on single and multiple track. Other objectives included determining the influence of crossing angle, train speed ratio, and train speed difference on the effectiveness of warning devices. Overall results confirmed effectiveness values developed previously (but with smaller data bases) for upgrades from passive systems to flashing lights (69 percent) and from passive systems to gates (84 percent). The only marked change from previous studies occurred in the flashing-lights-to-gates category; the effectiveness value determined in this study (72 percent) was higher than values obtained in previous work. Upgrades of warning devices on single track had higher effectiveness values than those on multiple tracks. Variation in train speeds at grade crossings, as measured by the speed-ratio and speed-difference concepts, had no apparent influence on the effectiveness of warning devices. Additional detailed conclusions as well as recommendations for further study are also included in the paper.

Safety at rail-highway grade crossings has long been a concern of many communities and public and private organizations. Although railroad grade-crossing accidents account for less than 1 percent of all motor vehicle accidents nationwide, the ratio of persons killed and injured to the number of grade-crossing accidents is an order of magnitude higher than that of all motor vehicle accidents. Consequently, substantial sums of money are spent each year to install warning devices at rail-highway grade crossings.

Attempts to apply warning devices to reduce the number of accidents at rail-highway grade crossings have a long history, dating from the earliest days of motor vehicle travel. The recent emphasis on grade-crossing safety has focused on optimizing the use of the limited funds available for upgrading crossings. A resource allocation model to assist states and railroads in determining the most effective allocation of funds for rail-highway crossing safety improvements has recently been developed by the U.S. Department of Transportation (DOT). The model, which has been described by Farr and Tustin (1), determines which crossings should have warning devices installed so as to achieve the maximum crossing safety benefit for a given level of funding. A brief description of the model is presented in the following paragraphs.

The typical approach in decision making on improvement of crossing safety is first to rank all crossings under consideration by using a hazard model. The most hazardous crossings are selected from this list of candidates for further review. The final decision is based on information gathered from on-site visits, the applicability of available alternatives, and the expected safety improvement. The resource allocation model includes a quantitative measure of safety benefit and equipment installation cost, along with a hazard value. Instead of providing a list of the most hazardous crossings, the model provides a list of the most cost-effective improvement decisions. These decisions are then examined and either adopted or rejected based on site-specific information.

The model is designed to rank crossings in the order that they need improvement and to recommend the warning device that should be installed to be the most cost and safety effective. Inputs to the resource allocation model include the predicted accident rates of the crossings, costs and effectiveness values of the different safety improvement options (such as flashing lights and gates), and the budget level available for safety improvement. To support the resource allocation model, costs and effectiveness values of different safety improvements were developed by using national data.

Several aspects of the model suggest that additional research is needed. One aspect is the effectiveness of different types of grade-crossing improvements. Effectiveness of a warning device is defined as the fraction by which accidents are reduced after installation of the warning device. This issue is important not only because it affects allocation of scarce highway resources, but also because it could affect the legal liability of railroads and states due to choice of crossing protection at a particular location.

Until the development of the resource allocation model, most measures of effectiveness of using gates versus using flashing light signals had been based on a study performed by the California Public Utilities Commission (PUC) in the early 1970s (2). Measures of effectiveness were developed for three types of improvements: passive system to flashing lights, passive system to gates, and flashing lights to gates. However, the universal applicability of the California PUC results has been questioned. Morrissey (3) noted that the effectiveness values from the PUC study were frequently criticized as being too high in view of accident statistics published by the Federal Railroad Administration (FRA).

Morrissey (3) undertook a study to improve the quality of and confidence in the data required for the DOT resource allocation model by determining new effectiveness values by using national data. The new effectiveness values were based on an analysis of the accident history of about 50 percent of the crossings (2,994) in the United States that had warning device upgrades during the period January 1, 1975, to December 31, 1978. Necessary data for the analysis were obtained from the DOT-Association of American Railroads (AAR) National Rail-Highway Crossing Inventory and the FRA Railroad Accident/Incident Reporting System. Morrissey's (3) effectiveness values almost equaled the results of the PUC study; that is, the California PUC results were within the 95-percent confidence intervals of the results of Morrissey's study.

The close agreement between Morrissey's results and those of the California PUC is not surprising because his study essentially repeated the PUC study, although with a much larger and more current data base. However, several aspects of Morrissey's study suggested a need for additional research. Since then, several more years of data have become available, thereby expanding opportunities for determining the effectiveness of various warning devices.

To address some of the questions raised by Morrissey's study, FHWA conducted additional investigations into the effectiveness of warning devices at rail-highway crossings. In addition to examining the three warning device upgrade categories studied by the PUC and Morrissey, Farr and Hitz (4) also obtained effectiveness values for upgrades to illumination and to cantilevered and mast-mounted flashing lights. Furthermore, they determined the influence of number of highway lanes, number of tracks, and train speed on the effectiveness of warning devices. The new effectiveness values determined for flashing lights and gates revealed results that were different from those of Morrissey's study. However, the results were claimed to be more accurate because the larger sample size used resulted in smaller confidence intervals than resulted from the sample size used by Morrissey. Farr and Hitz (4) also found that the effectiveness of warning devices declined with increasing number of tracks for grade crossings with two highway lanes. In general, train speed did not influence the effectiveness of warning devices.

Currently there are a number of issues concerning the effectiveness of warning devices that need to be addressed. Questions might be raised about the effectiveness of gates versus flashing lights at locations where warrants for gates are not met (e.g., a single track crossing with low to moderate train speeds). This is an important question because it affects the resource allocation model, and thereby influences the legal liability of railroads and states because of a choice of flashing lights rather than flashing lights and gates at a particular location. Additional efforts to stratify the data further to develop measures of effectiveness for the installation or upgrading of devices under various circumstances are definitely needed and have been noted by Farr and Hitz (4). Estimates of the effectiveness of stop signs would be desirable. This is a standard highway sign that may have a level of effectiveness that is greater than crossbucks. Several other potentially important factors should be analyzed to determine their influence on the effectiveness of warning devices. These factors include crossing angle and the ratio of maximum timetable speed to actual train speeds. For the latter factor, a crossing with a high timetable speed but a predominance of slow-speed trains that is not protected by a constant-warning-time device may create problems of credibility with motorists.

STUDY OBJECTIVES

The overall objective of the proposed research was to analyze further the national DOT-AAR inventory and FRA accident files to develop measures of effectiveness for the installation or upgrading of rail-highway grade-crossing protection devices under various conditions. Specific objectives of the study were

1. To develop measures of effectiveness for the following crossing upgrade stratifications: (a) passive warning device to flashing lights (single track), (b) flashing lights to gates (single track) either due to accidents or high train speeds ($\geq$ 50 mph), (c) flashing lights to gates (multiple track), (d) passive warning device to gates (single track), and (e) passive warning device to gates (multiple track);
2. To determine the influence of angle of crossing on the effectiveness of warning devices; and
3. To determine the influence of speed ratio (ratio of maximum timetable speed to typical minimum speed) and speed difference (difference between maximum timetable speed and typical minimum speed) on the effectiveness of warning devices for upgrades from (a) passive warning devices to flashing lights, (b) passive warning devices to crossing gates, and (c) flashing lights to crossing gates.

DATA ANALYSIS

The DOT-AAR Crossing Inventory File and the FRA Accident Data File for the period January 1, 1975, through December 31, 1982, were obtained from FRA. Of particular interest in this study was the classification of warning devices. The inventory file assigned a warning device class to each grade crossing. The FRA classes include eight categories of warning devices that reflect the level of motorist warning present. In general, the higher the class, the more warning information is provided to the motorist.

The first four warning device classes (no signs, other signs, stop signs, and crossbucks) are referred to as passive devices. Classes 5, 6, and 7 (special devices, wigwags or bells, and flashing lights, respectively) have usually been grouped into the flashing-light category (active devices). However, because classes 5 and 6 are infrequently used and often do not meet appropriate traffic engineering guidelines, these two classes were deleted from the flashing-light category in this study to provide more meaningful results. Class 8 of warning devices (flashing lights with gates) represents the most extensive type of crossing protection.

The data set that was created after working with the inventory data base included 13,852 warning device changes at public grade crossings. This data set was then merged with the accident file data base. To determine the effectiveness value for each upgrade category, or warning device, the average accident rates (accidents per crossing year) for crossings before and after installation of warning devices were compared.

The following formula (3) was used to calculate the effectiveness of the warning devices:

$$E = (A_b/Y_b - A_a/Y_a)/(A_b/Y_b) \tag{1}$$

where

E = effectiveness of a particular warning device;

TABLE 1 Summary of Results of Effectiveness Values for Flashing Lights and Gate Upgrades

Upgrade Category[a]	No. of Crossings	Before Upgrade Accidents	Before Upgrade Crossing Years	After Upgrade Accidents	After Upgrade Crossing Years	Effectiveness Value (%)	Standard Deviation of Effectiveness Value (%)	95 Percent Confidence Interval (%)
P to FL	2,786	1,407	10,824	448	11,234	69	1.6	66-72
P to G	2,781	2,157	10,934	352	11,291	84	0.9	82-86
FL to G	2,167	2,139	8,179	639	8,838	72	1.2	70-75
Total	7,734	5,703	29,937	1,439	31,363	76	0.7	74-77

[a]P = passive, FL = flashing lights, and G = flashing lights with gates.

A_b = total number of accidents before warning device installation;
Y_b = total number of cross years before warning device installation;
A_a = total number of accidents after warning device installation; and
Y_a = total number of crossing years after warning device installation.

Results of the computations of effectiveness values are presented in the following section.

RESULTS

Overall Effectiveness

Currently, the rail-highway crossing resource allocation model considers three categories of warning device upgrades: (a) passive systems to flashing lights, (b) passive systems to gates, and (c) flashing lights to gates. Effectiveness values and confidence intervals for upgrades within these warning device categories were calculated and compared with similar results from earlier studies. This was done both to serve as a check on the methodology used herein and to examine whether there would be any changes in effectiveness values with the larger sample size used in this study. Results are presented in Table 1. As the data in Table 2 indicate, the results are slightly different from those obtained in previous studies. Because of the additional data used in the current study, the results are more accurate, as indicated by the smaller confidence intervals.

A review of the data in Table 2 indicates that only effectiveness values for flashing-lights-to-gate upgrades have changed markedly from previous studies. Farr and Hitz (4) had noticed a similar phenomenon and noted that it was difficult to explain. They hypothesized that flashing-light crossings more recently selected for upgrading to gates had unique characteristics that caused gates to be particularly effective relative to flashing lights.

It is also possible that the increased effectiveness is because of improved traffic engineering (as it applies to the layout of the displays) at crossings that have been recently upgraded. When flashing lights are upgraded to flashing lights and gates, an entirely new crossing installation, including both displays and control circuitry, generally results. Further, since the completion of the California PUC study (2), motion sensors and predictors have come into common use. This increases the credibility of the device and may contribute to the higher effectiveness value.

Single-Track Upgrades

Three different stratifications of warning device upgrades on single tracks were examined: (a) passive systems to flashing lights, (b) passive systems to gates, and (c) flashing lights to gates. The data in Table 3 present the effectiveness values and confidence intervals for each of these upgrade categories. Unexpectedly, warning device upgrades on single tracks had a higher effectiveness value than those on multiple tracks.

As anticipated, the highest effectiveness value (86 percent) was associated with upgrades from passive devices to flashing lights with gates. Upgrades from flashing lights to gates had an effectiveness value of 74 percent. The lowest effectiveness value of the three upgrades was associated with the passive-to-flashing-lights condition (71 percent). Because of the large sample size involved, confidence intervals were of approximately the same width as those for the overall analysis presented previously.

As a subset of the analysis just described, the upgrading of warning devices from flashing lights to gates on single track under the circumstances of (a) accidents and (b) high train speeds (maximum timetable speed of 50 mph or greater) was examined. The data in Table 4 present effectiveness values for those crossings that experienced accidents before the upgrade occurred. Also presented in Table 4 are effectiveness values for crossings that experienced one or more accidents either before the upgrade or after. Crossings that did not experience accidents before or after the upgrade were excluded because it was thought that such crossings would not aid in the

TABLE 2 Comparison of Effectiveness Values for Flashing Lights and Gate Upgrades for Current and Previous Studies

Upgrade Category[a]	Effectiveness Values: Current Study	Farr and Hitz (1982)	Morrissey (1981)	California PUC (1974)	95 Percent Confidence Interval: Current Study	Farr and Hitz (1982)	Morrissey (1981)	California PUC (1974)
P to FL	69	71	65	64	66-72	66-75	57-73	NA
P to G	84	82	84	88	82-86	79-85	80-89	NA
FL to G	72	69	64	66	70-75	65-73	56-71	NA

Note: NA = not available.
[a]P = passive, FL = flashing lights, and G = flashing lights with gates.

TABLE 3 Summary of Results of Effectiveness Values for Single- and Multiple-Track Upgrades

Upgrade Category[a]	No. of Crossings	Before Upgrade Accidents	Before Upgrade Crossing Years	After Upgrade Accidents	After Upgrade Crossing Years	Effectiveness Value (%)	Standard Deviation of Effectiveness Value (%)	95 Percent Confidence Interval (%)
Single Track								
P to FL	2,488	1,287	9,686	390	10,079	71	1.63	68–74
P to G	2,089	1,584	8,110	234	8,609	86	0.95	84–88
FL to G	1,626	1,539	6,187	437	6,703	74	1.34	71–76
Multiple Track								
P to G	567	520	2,311	113	2,202	77	2.27	73–81
FL to G	483	569	1,961	193	1,894	65	2.70	60–70

[a]P = passive, FL = flashing lights, and G = flashing lights with gates.

TABLE 4 Effectiveness Values for Upgrades from Flashing Lights to Gates on Single Tracks

Maximum Train Speed (mph)	No. of Crossings	Before Upgrade Accidents	Before Upgrade Crossing Years	After Upgrade Accidents	After Upgrade Crossing Years	Effectiveness Value (%)	95 Percent Confidence Interval (%)
Crossings with Accidents Before							
0–49	456	1,054	2,137	140	1,671	83	80–86
≥ 50	231	475	1,062	80	959	81	77–85
Crossings with or without Accidents Before							
0–49	1,007	1,054	4,062	253	4,000	76	72–79
≥ 50	619	475	2,252	184	2,705	68	63–73

examination of high accident experience in which train speeds were 50 mph or greater. Although the effectiveness values for the 0-to-49-mph category were slightly higher than values for the high-speed category (83 and 81 percent, respectively), the difference was not significant. This result was unexpected because it was thought that installation of gates would be more effective at crossings where high train speeds are encountered.

When the crossings were added to the analysis that did not experience accidents before the upgrade but did have accidents after the upgrade, the effectiveness values for the two speed categories were significantly different. As expected, the effectiveness values were lower than for crossings that experienced accidents before the upgrade. The 0-to-49-mph category again revealed a higher effectiveness value (76 versus 68 percent) than did the high-speed category. Once again, this is contrary to the notion that effectiveness of gates should be higher at crossings with high-speed trains. One possible explanation for these results is the credibility problem created when the majority of rail traffic on a line travels at speeds substantially less than the timetable speed. This problem will be discussed in a subsequent section.

Multiple-Track Upgrades

Two different stratifications of warning device upgrades on multiple tracks were examined: (a) passive systems to gates and (b) flashing lights to gates. Effectiveness values and confidence intervals for each of these categories are given in Table 3. Note that these upgrades had a lower effectiveness value than did the corresponding upgrades on single track.

As expected, upgrades from passive devices to flashing lights had a higher effectiveness value (77 percent) than did upgrades from flashing lights to gates (65 percent). In both cases the confidence intervals were larger than those obtained in the overall or single-track analyses because of the relatively small sample sizes involved.

It is hypothesized that the lower effectiveness value associated with multiple-track upgrades is due to the greater exposure (product of train times vehicular volumes) likely to be found at these crossings. Accidents continue to occur because of high exposure levels even after gates have been installed at the crossing. Ideally, the accident rates used in developing the effectiveness values should include a measure of exposure.

A further investigation involving improvement types and before-accident rates for active and passive devices was made to determine if any additional conclusions could be drawn. Before-accident rates in terms of accidents per crossing year were computed; these are given in Table 5. Note the relatively high rates in the flashing-lights-to-gates category for both single and multiple track. It can be inferred from the data that flashing lights are upgraded to gates in response to an accident problem. Although this is good management, it does tend to bias the effectiveness data.

TABLE 5 Before-Accident Rates for Single- and Multiple-Track Upgrades

Upgrade Category[a]	Accidents per Crossing Year: Single Track	Accidents per Crossing Year: Multiple Track
P to FL	0.133	–
P to G	0.195	0.225
FL to G	0.250	0.290

[a]P = passive, FL = flashing lights, and G = flashing lights with gates.

Influence of Crossing Angle

The effectiveness of crossing warning devices would be expected to be related to the angle of the crossing. Effectiveness of devices should be greatest at oblique-angle crossings because it is at these locations that motorists may have difficulty determining the exact location of the track. They may also have trouble detecting an approaching train because of sight obstructions in the vehicle or because of uncertainty in determining where to look along the tracks.

In this part of the research, crossing-angle categories of 0 to 30 degrees, 30 to 60 degrees, and 60 to 90 degrees were analyzed for their influence on the effectiveness of warning devices. The influence of angle of crossing on the effectiveness of warning devices is given in Table 6 for both single- and multiple-track categories.

For the single-track condition, a review of the data in Table 6 indicates that for upgrade categories of passive to flashing lights and flashing lights to gates, the effectiveness values are greatest in the angle-of-crossing category of 60 to 90 degrees. This is contrary to the hypothesis stated at the beginning of this section. As expected, the effectiveness of single-track upgrades from passive devices to gates was greatest in the oblique-angle categories, with an 88-percent effectiveness. Note that the confidence intervals were rather wide because of the relatively small sample sizes in each of the upgrade categories.

Results for multiple-track crossings did not reveal a definite pattern like the single-track crossings. Highest effectiveness value (83 percent) was for the passive-to-gates upgrade with a crossing angle of 0 to 29 degrees. This outcome was expected. However, it was not expected that, for this same type of upgrade, the angle-of-crossing category of 30 to 59 degrees would be associated with the lowest effectiveness value (70 percent). The opposite results were obtained for the flashing-lights-to-gates upgrades. In this case, the 30-to-59-degree category had the highest effectiveness value (70 percent). The 0-to-29-degree category was associated with the lowest effectiveness value (63 percent). Because of the small number of crossings in the multiple-track categories, some of the effectiveness values had large confidence intervals.

Although the lack of any definite pattern as far as variation in effectiveness with angle of crossing was unexpected, the results may be explained by two items of information that are not included in the data base. The first of these is sight distance. Sight distance is not quantified in the data base nor are sight obstructions (such as vegetation or structures) noted. The second factor is the direction of approach of vehicular and train traffic. Both of these factors are important in determining when drivers first detect the presence of trains, yet neither is included in the data base.

Influence of Train Speed

At a number of crossings, activation of warning devices is based on the maximum timetable speed of trains. However, the majority of rail traffic on the line may travel at speeds substantially slower than the timetable speed. This creates credibility problems with motorists in that some drivers may try to proceed past flashing lights or maneuver around crossing gates because of the lengthy time interval between signal activation and actual passage of the train through the crossing. The influence of train speed on the effectiveness of warning devices was examined by using two different measures. One concept was the speed-difference approach, in which speed difference was calculated as the algebraic difference between maximum timetable speed and typical minimum speed. Also, a speed-ratio concept was examined, in which the ratio of maximum timetable speed to typical minimum speed was computed for crossings. Influence of train speed on the effectiveness of warning devices was determined for the three categories of flashing lights and gate upgrades.

Figures 1 and 2 show that there was no relationship between train speed difference and effectiveness values in any upgrade category for single-track and multiple-track crossings, respectively. Spearman's Rho statistical tests for trend were performed; they indicated that there were no trends (either upward or downward) at the 95-percent confidence level.

Figures 3 and 4 show the plots of effectiveness values versus train speed ratio. Spearman's Rho tests revealed no significant trends in any of the

TABLE 6 Influence of Angle of Crossing on Effectiveness of Warning Device

Upgrade Category[a]	Angle of Crossing (degree)	No. of Crossings	Before Upgrade: Accidents	Before Upgrade: Crossing Years	After Upgrade: Accidents	After Upgrade: Crossing Years	Effectiveness Value (%)	95 Percent Confidence Interval (%)
Single Track								
P to FL	0–29	246	160	990	52	983	67	57–77
	30–59	436	177	1,739	60	1,776	67	57–76
	60–90	1,774	950	6,948	278	7,313	72	69–76
P to G	0–29	252	208	898	31	1,127	88	84–92
	30–59	266	194	1,048	25	1,095	88	83–93
	60–90	1,550	1,182	6,165	178	6,386	85	83–88
FL to G	0–29	119	158	678	59	749	66	57–76
	30–59	225	157	818	65	975	65	56–75
	60–90	1,221	2,214	4,691	313	4,971	76	73–79
Multiple Track								
P to G	0–29	66	91	288	13	238	83	73–92
	30–59	73	59	303	16	278	70	55–86
	60–90	428	370	1,719	84	1,685	77	72–82
FL to G	0–29	39	59	159	21	153	63	47–79
	30–59	60	64	258	17	227	70	55–85
	60–90	384	446	1,543	155	1,514	65	59–71

[a]P = passive, FL = flashing lights, and G = flashing lights with gates.

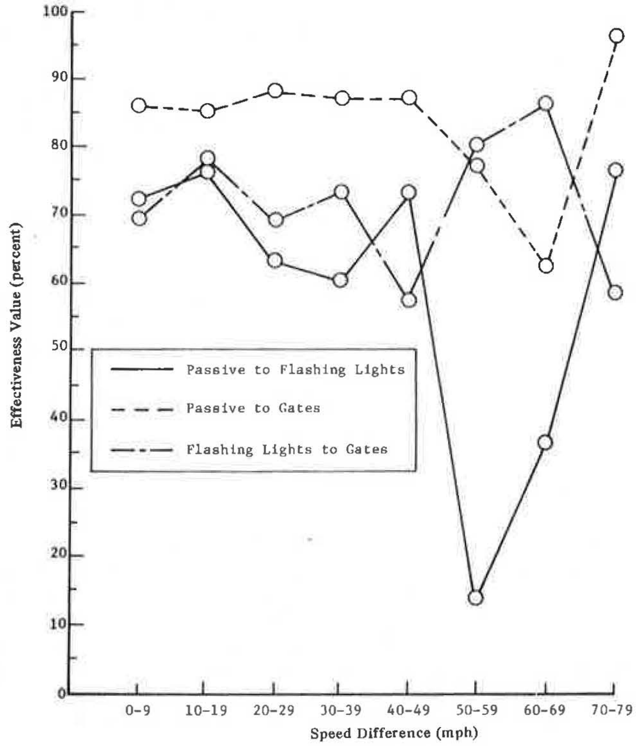


FIGURE 1 Relationship between speed difference and effectiveness values for flashing lights and gate upgrades at single-track crossings.

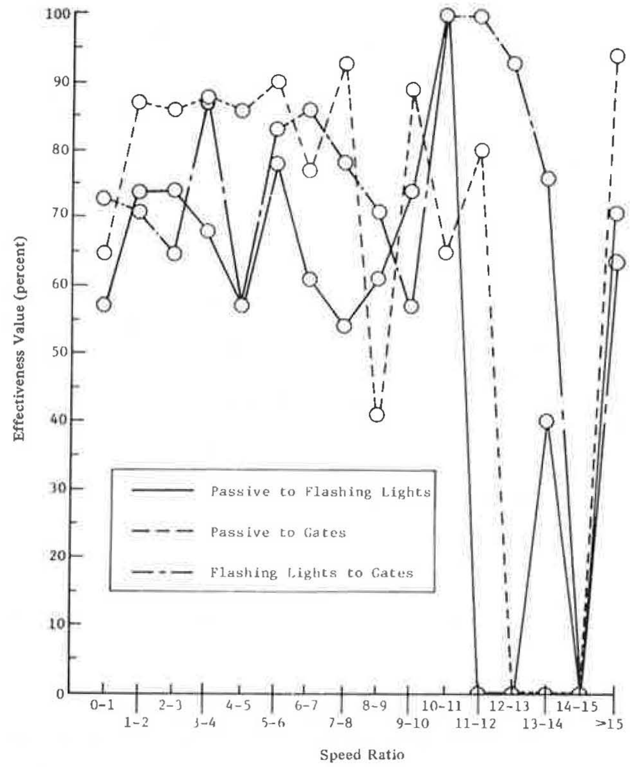


FIGURE 3 Relationship between speed ratio and effectiveness values for flashing lights and gate upgrades at single-track crossings.

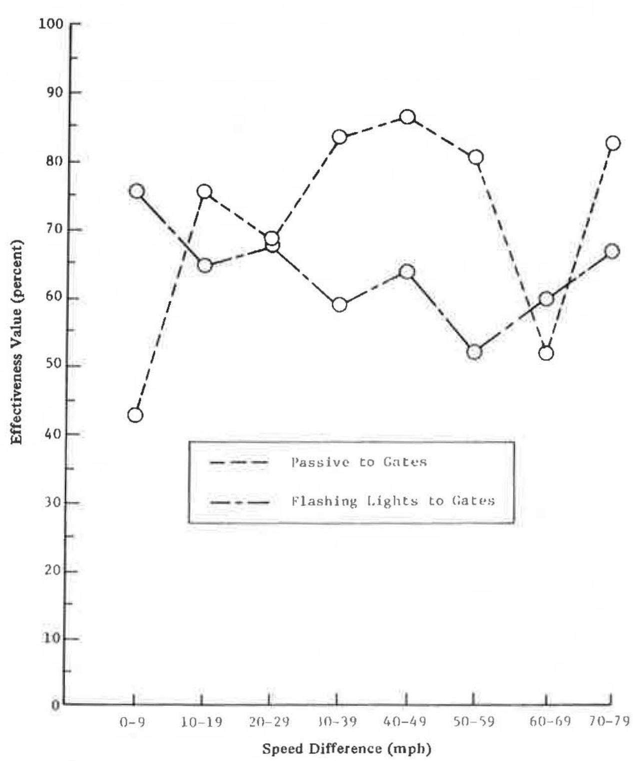


FIGURE 2 Relationship between speed difference and effectiveness values for flashing lights and gate upgrades at multiple-track crossings.

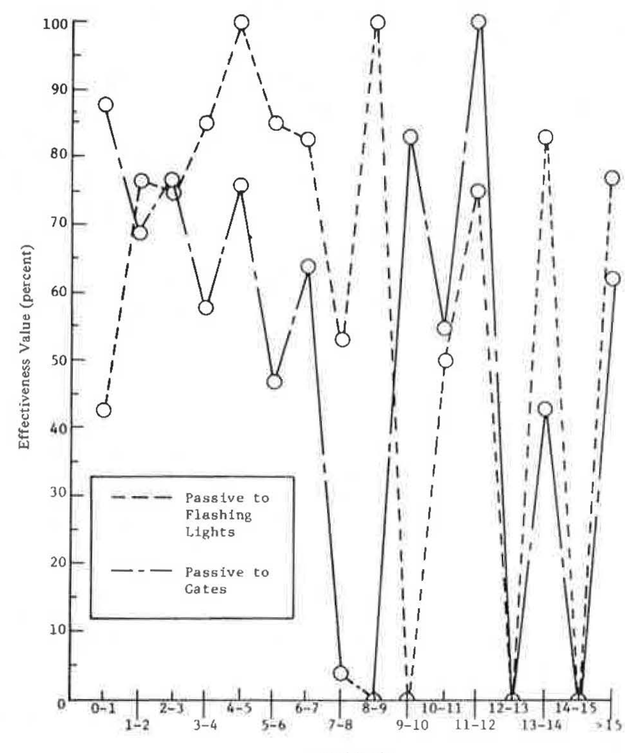


FIGURE 4 Relationship between speed ratio and effectiveness factors for flashing lights and gate upgrades at multiple-track crossings.

upgrade categories. Thus it was concluded that train speed ratio had no influence on the effectiveness of warning devices.

The results of these analyses were unexpected. Train speed, as measured by the two concepts used here, has no apparent influence on the effectiveness of warning devices for flashing lights and gate upgrades.

CONCLUSIONS

Results of this study generally confirmed the effectiveness values that were developed previously for upgrades from passive systems to flashing lights and from passive systems to gates. Upgrades from passive systems to flashing lights had an effectiveness value of 69 percent, whereas upgrades from passive systems to gates had an effectiveness value of 84 percent. The only marked change from results of previous studies occurred in the flashing-lights-to-gates upgrade category. The effectiveness value determined in this study was 72 percent. This is higher than values obtained in previous studies. There is no readily available explanation for this phenomenon.

Other important conclusions drawn from the study are as follows:

1. Warning device upgrades on single track had higher effectiveness values than those on multiple track. In both cases the highest effectiveness value was associated with upgrades from passive devices to flashing lights with gates.
2. There was no significant difference in effectiveness values between the 0-to-49-mph and the 50-mph-and-greater speed categories, for upgrades from flashing lights to gates at single-track crossings that had accidents before the upgrade.
3. Flashing lights appear to be upgraded to gates in response to an accident problem. Although this is good management, it tends to bias the effectiveness data.
4. Effectiveness of upgrades from passive devices to gates at single-track crossings was greatest in the oblique-angle categories (88-percent effectiveness). However, for passive-to-flashing-lights and flashing-lights-to-gates upgrades, effectiveness values were greatest in the 60-to-90-degree-angle category. Results for multiple-track crossings failed to show a definite pattern.
5. Variation in train speeds at crossings, as measured by the speed-difference and speed-ratio concepts, had no apparent influence on the effectiveness of warning devices for flashing lights and gate upgrades.

RECOMMENDATIONS

Results of this study suggest a number of areas in which additional research could prove fruitful. These are outlined in the following paragraphs.

Reference was made in a preceding section to the motorist credibility problem that exists at certain grade crossings that are equipped with active warning devices. There are two basic types of control systems for active devices: (a) fixed-distance concept and (b) constant-warning-time concept. With fixed-distance systems, trains activate the signals or gates a predetermined distance from the crossing. The major drawback to such systems is that warning devices operate continuously while the train is on the approach track circuit, regardless of train speed. Motorists may become impatient in situations in which the warning device is active for a long time (e.g., slow train speed). 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 device. Thus, regardless of train speed, a uniform warning time is provided. With constant-warning-time systems, trains can move or switch on the approaches without reaching the crossing and, depending on their speed, never cause the crossing warning devices to be activated.

It could be hypothesized that the greater the difference between typical train speeds and maximum timetable speed (the basis on which the signals were designed), the higher the accident rate at crossings equipped with fixed-distance systems. Additional research is warranted to analyze the DOT-AAR data files to determine if accident frequency and characteristics at crossings with fixed-distance systems differ from those crossings with constant-warning-time systems. Note, however, that the results will have little meaning unless the accident data are normalized for exposure (traffic volume times train volume).

The need to develop a means to normalize exposure data is critical. This affects the analysis of constant-warning-time devices and may also account for the lower effectiveness of cantilever flashers noted by Farr and Hitz (4). Constant-warning-time devices tend to be used on more important rail lines on which there are more train movements and thus higher exposure. Similarly, cantilever flashers are frequently used at crossings on multilane highways, at which vehicular exposure would be higher.

Results of this study indicated no definite relationship between angle of crossing and effectiveness values. It was hypothesized that this was due to lack of information about sight distances and directions of approach of vehicular and train traffic at the grade crossing. Further study is warranted, perhaps involving field investigations, to determine the influence of these variables on accident experience. Development of a simple yet meaningful way to incorporate such factors into either the inventory or accident data base appears to be appropriate.

Another fruitful area of future research would be the development of capital and life-cycle costs for each of the upgrade categories studied in this project. These are needed if the results of this project are to be applied to the resource allocation model. This information can also be used to determine the relative cost-effectiveness of various improvement alternatives (such as flashing lights versus lights with gates on a single-track crossing).

There is one other area, closely related to the determination of effectiveness values, in which additional research would be desirable. In this study, confidence intervals were calculated by using a relationship developed by Morrissey (3). However, calculation of appropriate confidence intervals for effectiveness factors is subject to some interpretation because of the unique statistical nature of crossing accidents. Statisticians contacted by the investigators pointed out the need for a more thorough derivation of confidence interval formulas for the effectiveness studies. Determining the variance of ratios, such as the accident rates considered here, was beyond the scope of this project because of its theoretical complexity. Additional research of a statistical nature is needed to determine whether the confidence interval used here and in previous studies is appropriate and, if not, to develop a true confidence interval.

ACKNOWLEDGMENTS

This paper is based on research sponsored by the West Virginia Department of Highways (WVDOH) in cooperation with FHWA, U.S. Department of Transportation.

Special appreciation is expressed to Ray Lewis (WVDOH) for his feedback on project activities, to Bruce George (FRA) for furnishing the data and providing helpful advice during preliminary data analysis, and to Edwin Farr and John Hitz (Transportation Systems Center) for their assistance and encouragement throughout the project.

REFERENCES

1. E. Farr and B.H. Tustin. Optimizing Resources at Rail-Highway Crossings. ITE Journal, Vol. 52, No. 1, Jan. 1982, pp. 25-28.
2. The Effectiveness of Automatic Protection in Reducing Accident Frequency and Severity at Public Grade Crossings in California. Railroad Operations and Safety Branch, California Public Utilities Commission, San Francisco, June 30, 1974, 196 pp.
3. J. Morrissey. The Effectiveness of Flashing Lights and Flashing Lights and 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, 20 pp.
4. E. Farr and J. Hitz. Additional Investigations into Rail-Highway Crossing Warning Device Effectiveness. Draft Report. FHWA, U.S. Department of Transportation, April 1982, 32 pp.

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

Effectiveness of Constant-Warning-Time Versus Fixed-Distance Warning Systems at Rail-Highway Grade Crossings

JOHN A. HALKIAS and RONALD W. ECK

ABSTRACT

The study objective was to determine the influence of road classification, angle of crossing, and train speed on the effectiveness of fixed-distance and constant-warning-time systems at public rail-highway grade crossings. Data were acquired from the U.S. Department of Transportation-Association of American Railroads Crossing Inventory File and the FRA Accident/Incident Reporting System for the period January 1, 1975, through December 31, 1982. Fixed-distance and constant-warning-time systems revealed similar effectiveness values (82 and 85 percent, respectively) when changed from passive devices. For changes from fixed-distance to constant-warning-time systems, the effectiveness value was 26 percent. This result tended to confirm the hypothesis that constant-warning-time systems have greater credibility with motorists than do fixed-distance systems. Functional class of road had no apparent influence on the effectiveness of warning systems for upgrades to fixed-distance systems and constant-warning-time systems. The effectiveness of upgrades in the fixed-distance-to-constant-warning-time class was greatest for the angle-of-crossing category of 0 to 29 degrees (68 percent). For passive-to-fixed-distance and passive-to-constant-warning-time upgrades, effectiveness values in the 60-to-90-degree-angle category were essentially equal to those in the oblique-angle categories (82 percent). For constant-warning-time systems, effectiveness increased with increase in variation of train speed. Train speed, as measured by the concepts of speed ratio and speed difference, had no apparent influence on warning system effectiveness for either system.

The Federal Railroad Safety Act of 1970 and the Federal Highway Safety acts of 1970 and 1973 required the Secretary of Transportation to take action to improve rail-highway grade crossing safety. In response to these mandates, the National Rail-Highway Crossing Inventory and the Railroad Accident/Incident Reporting System were implemented (1). These data bases, which are updated on a regular basis, are used extensively by federal, state, and railroad company planners and decision makers as well as by researchers. The files are important inputs to the U.S. Department of Transportation (DOT) resource allocation procedure and accident prediction equations for rail-highway grade crossings. The generally declining trend in rail-highway grade-crossing fatalities since the mid-1970s can probably be at least partially attributed to the improved decision making made possible by these data bases.

To support resource allocation decisions, costs and effectiveness values of different safety improvements are needed (2) and have been developed by using national data. The effectiveness values used represent the percentage reduction in accidents expected from installation of a particular type of warning device at a typical crossing. Currently (3), effectiveness values are required for the three types of warning device installations considered by the DOT resource allocation procedure: (a) flashing lights installed at passively signed crossings, (b) gates installed at passively signed crossings, and (c) gates installed at crossings with flashing lights.

The existence of the FRA data bases has prompted a number of recent research efforts (4,5) to develop effectiveness values for other types of warning device installations for possible consideration by the resource allocation procedure. The authors (6) developed measures of effectiveness for warning devices under a variety of conditions. It was found that variation in train speed had no apparent influence on warning device effectiveness for flashing lights and gate upgrades. This result was unexpected, as will be discussed later.

There are two basic types of control systems for active (i.e., flashing lights or gates) warning devices: (a) fixed-distance concept and (b) constant-warning-time concept. With fixed-distance systems, trains activate the signals or gates at a predetermined distance from the crossing. This distance is calculated by using the speed of the fastest train and a specified minimum warning time. The major drawback to such systems is that warning devices operate continuously while the train is on the approach track circuit, regardless of train speed. Motorists may become impatient in situations in which the warning device is active for a long time (e.g., slow train speed). This creates credibility problems with motorists in that some drivers may try to proceed past flashing lights or maneuver around crossing gates because of the lengthy time interval between signal activation and actual passage of the train through the crossing.

Constant-warning-time systems provide the most desirable type of train detection at crossings where trains traveling at widely different speeds use the crossing. 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 device. Thus, regardless of train speed, a uniform warning time is provided.

It could be hypothesized that the greater the difference between typical train speeds and maximum timetable speed (the basis on which signals are typically designed), the greater the effectiveness of devices upgraded to constant-warning-time systems. The results just mentioned (6), from situations in which variation in train speeds had no influence on the effectiveness of warning devices, included the aggregate of fixed-distance and constant-warning-time systems; no detailed breakdown was available from the FRA data base. Additional study appeared warranted to analyze the DOT data base to determine if the effectiveness of warning devices at crossings with fixed-distance systems differs from that at crossings with constant-warning-time systems.

Farr and Hitz (3) investigated the effectiveness of constant-warning-time devices. Two crossing upgrade categories were examined: (a) flashing lights without constant warning time upgraded to flashing lights with constant warning time, and (b) gates without constant warning time upgraded to gates with constant warning time. The results were unsatisfactory because there were only 39 upgrades in the first category (117.6 crossing years of data before upgrade and 113.2 crossing years of data after upgrade) and 80 upgrades in the second category (213.4 crossing years before and 259.9 crossing years after). The confidence intervals were too large to provide any meaningful estimates of effectiveness. Further investigation of this issue using additional data available in the inventory and accident files appears appropriate.

The overall goal of the study described here was to develop measures of effectiveness for fixed-distance and constant-warning-time systems under several conditions. Specific objectives of the study were to determine the influence of each of the following variables on the effectiveness of the two different warning systems:

1. Road classification,
2. Angle of crossing, and
3. Train speed, in particular, speed difference, speed ratio, and maximum speed.

DATA ANALYSIS

The DOT-Association of American Railroads (AAR) Crossing Inventory File and the FRA Accident Data File for the period of January 1, 1975, through December 31, 1982, were obtained from FRA. Of particular interest in this study was the classification of warning devices. The inventory file assigns a warning device class to each grade crossing. The FRA classes include eight categories of warning devices that reflect the level of motorist warning present. In general, the higher the class, the more warning information provided to the motorist.

The first four warning device classes (no signs, other signs, stop signs, and crossbucks) are referred to as passive devices. Classes 5, 6, and 7 (special devices, wigwags or bells, and flashing lights, respectively) have usually been grouped into the flashing-light category (active devices). However, because classes 5 and 6 are infrequently used and often do not meet appropriate traffic engineering guidelines, these two classes were deleted from the flashing-light category in this study to provide more meaningful results. Class 8 of the warning devices (flashing lights with gates) represent the highest type of crossing warning. The existence of constant-warning-time systems at crossings with active devices was indicated in the inventory as a positive response to the question: Does crossing signal provide speed selection for trains?

The data set that was created for the overall effectiveness of fixed-distance and constant-warning-time systems, after working with the inventory data base, included 3,195 warning device changes at public grade crossings. These 3,195 changes (Table 1) are warning device changes to fixed-distance and

TABLE 1 Number of Warning Device Change Records by System Type for Active Devices (1975-1982)

	Warning Device System Type After Change				
	Fixed Distance		Constant Warning Time		
Warning Device System Type Before Change	Flashing Lights	Gates	Flashing Lights	Gates	Total
No signs, stop signs, other signs, or crossbucks	1,061	1,103	192	405	2,761
Flashing lights with fixed distance	NC	NC	68	222	290
Flashing lights with constant warning time	19	0	NC	NC	19
Gates with fixed distance	NC	NC	0	122	122
Gates with constant warning time	0	3	NC	NC	3
Total	1,080	1,106	260	749	3,195

Note: NC indicates not considered in this study; these are changes from one warning device system to the same warning device system.

constant-warning-time systems (for both flashing lights and gates). Note that the number of warning device changes is substantially larger than that found by Farr and Hitz (3). This is believed to be due to the larger data base used in the current study and the fact that all warning device changes were examined, whereas Farr and Hitz (3) examined only the most recent change.

The inventory data base described in the preceding paragraph was then merged with the accident file data base. To determine the effectiveness value for each upgrade category, or warning device system, the average accident rates (accidents per crossing year) for populations of crossings before and after installation of warning devices were compared. The following formula (4) was used to calculate the effectiveness of the warning device systems:

$$E = (A_b/Y_b - A_a/Y_a)/(A_b/Y_b) \quad (1)$$

where

E = effectiveness of a particular warning device system,

A_b = total number of accidents before warning device installation,

Y_b = total number of crossing years before warning device installation,

A_a = total number of accidents after warning device installation, and

Y_a = total number of crossing years after warning device installation.

Results of the computations of effectiveness values are presented in the following section. In reviewing the data it should be noted that the FRA accident data base, which was used in compiling effectiveness values, has not been independently verified and represents only reflections of accident data as reported by railroad carriers.

RESULTS

Overall Effectiveness

Three main categories of warning system upgrades were considered in this study: (a) passive to fixed-distance system, (b) passive to constant-warning-time systems, and (c) fixed-distance to constant-warning-time systems. Effectiveness values and confidence intervals for upgrades within these categories were calculated and, where appropriate, compared with general results from earlier studies. Note that "downgrades" from constant-warning-time to fixed-distance systems were investigated initially, but this category had to be eliminated from further consideration because of its extremely small sample size. Overall results of effectiveness values are presented in Table 2. To provide additional insight,

TABLE 2 Summary of Results of Effectiveness Values for Upgrades to Fixed-Distance and Constant-Warning-Time Systems

		Before Upgrade		After Upgrade			
Upgrade Category	No. of Crossings	Accidents	Crossing Years	Accidents	Crossing Years	Effectiveness Value (%)	95 Percent Confidence Interval (%)
Passive to fixed distance							
Lights	1,061	449	3,715	156	4,706	73	68 to 78
Gates	1,103	802	4,310	102	4,506	88	86 to 90
Total	2,164	1,251	8,025	258	9,212	82	81 to 83
Passive to constant warning time							
Lights	192	80	551	31	925	77	68 to 86
Gates	405	266	1,395	44	1,785	87	83 to 91
Total	597	346	1,946	75	2,710	85	81 to 89
Fixed distance to constant warning time							
Lights	68	34	167	54	331	20	-11 to 51
Gates	122	23	258	39	608	28	-8 to 64
Total	190	57	425	93	939	26	3 to 49
Lights to gates	222	122	578	49	1,028	77	59 to 95

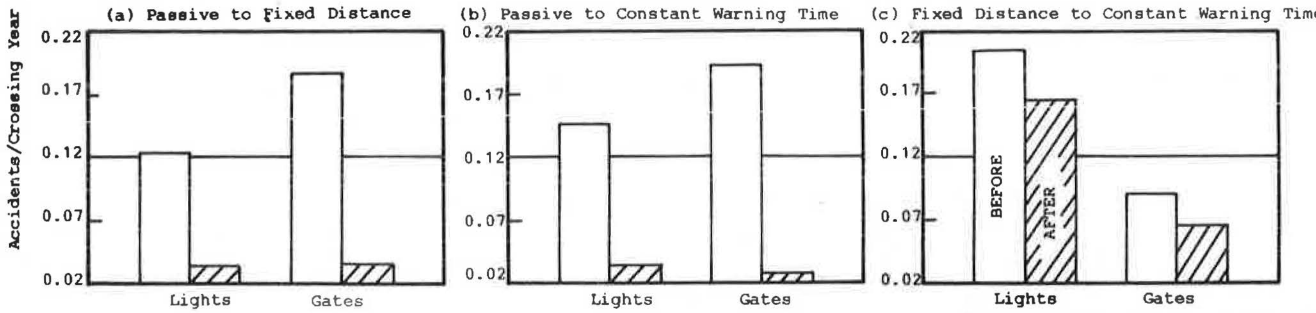


FIGURE 1 Graphical comparison of before-and-after accident rates for upgrades to fixed-distance and constant-warning-time systems.

graphical comparisons were made of the before-and-after accident rates on which the effectiveness values were calculated; these are shown in Figure 1.

Both fixed-distance and constant-warning-time systems revealed high effectiveness values (82 and 85 percent, respectively) when changed from passive devices. This is due in part to the change from use of passive devices to use of active warning devices. It was expected that these values would be high whether the system was a fixed-distance system or a constant-warning-time system. It was not expected that both upgrades would have essentially the same effectiveness value.

For changes from fixed-distance to constant-warning-time systems, two principal types of upgrades were considered: (a) flashing lights (or gates) of fixed distance to flashing lights (or gates) of constant warning time, and (b) flashing lights of fixed distance to gates of constant warning time. For the former case, the effectiveness was 26 percent. The 95-percent confidence interval, although rather wide because of the small sample size, did not include zero; thus there was a significant degree of effectiveness. Note that Farr and Hitz (3) had examined similar cases (but with smaller sample size) and found negative effectiveness values.

TABLE 3 Influence of Road Classification on Effectiveness of Active Warning Systems

Upgrade Category	No. of Crossings	Before Upgrade: Accidents	Before Upgrade: Crossing Years	After Upgrade: Accidents	After Upgrade: Crossing Years	Effectiveness Value (%)	95 Percent Confidence Interval (%)
Rural							
Passive to fixed distance							
Principal arterial	20	3	58	1	103	81	39 to 100
Minor arterial	78	39	237	6	386	91	83 to 99
Major collector	327	125	1,096	45	1,520	74	65 to 83
Minor collector	277	108	905	25	1,302	84	77 to 91
Local	683	346	2,730	57	2,714	83	78 to 88
Passive to constant warning time							
Principal arterial	11	1	24	1	64	63	-39 to 100
Minor arterial	23	11	62	5	117	76	52 to 100
Major collector	65	25	202	6	301	84	70 to 98
Minor collector	68	29	220	2	318	95	88 to 100
Local	199	89	659	25	917	80	71 to 89
Fixed distance to constant warning time							
Principal arterial	12	2	30	6	61	-48	-275 to 100
Minor arterial	23	7	47	6	126	68	35 to 100
Major collector	38	12	74	8	192	74	52 to 96
Minor collector	34	3	106	2	156	55	-25 to 100
Local	39	7	88	8	194	48	-3 to 99
Urban							
Passive to fixed distance							
Freeway	4	3	15	1	18	72	0 to 100
Principal arterial	67	50	257	9	271	83	71 to 95
Minor arterial	177	167	651	37	759	81	75 to 87
Collector	155	138	612	23	616	83	76 to 90
Local	367	269	1,433	52	1,491	81	76 to 86
Passive to constant warning time							
Freeway	2	2	5	6	11	-36	-200 to 100
Principal arterial	18	16	65	3	76	84	65 to 100
Minor arterial	48	41	153	9	229	85	75 to 95
Collector	62	58	215	5	276	93	87 to 99
Local	101	74	340	13	463	87	80 to 94
Fixed distance to constant warning time							
Freeway	8	6	23	4	41	63	20 to 100
Principal arterial	71	58	173	41	329	63	50 to 76
Minor arterial	72	44	182	24	325	69	55 to 83
Collector	34	13	76	15	160	45	7 to 83
Local	79	24	203	23	386	50	22 to 78

The effectiveness of upgrades from flashing lights of fixed distance to gates of constant warning time was rather high (77 percent). This is actually a special case of the fixed-distance-to-constant-warning-time category. Much of the effectiveness is probably caused by the concurrent upgrade in warning device type from flashing lights to gates; previous work (6) has indicated that such upgrades have effectiveness values in the range of 70 to 75 percent.

Influence of Road Classification

The influence of road classification on warning system effectiveness was analyzed to determine whether certain roadway types demonstrated different warning system effectiveness values than others. Implicitly associated with each roadway functional type would be information about certain crossing characteristics such as average daily traffic and urban versus rural environment. For example, Farr and Hitz (3) noted that the greater visual confusion that confronts motorists in urban areas may be responsible for the significantly lower effectiveness values for certain categories of warning device upgrades at urban crossings. Results of this analysis, given in Table 3 and shown in Figure 2, do not indicate any trends or significant differences in effectiveness values for either the urban or rural road classifications. Some effectiveness values are higher than others, but because of the large confidence intervals attributable to the relatively small sample sizes, no significant differences were noted.

Influence of Crossing Angle

The effectiveness of crossing warning devices would be expected to be related to the angle of the crossing. Device effectiveness should be greatest at oblique-angle crossings because it is at these locations that motorists otherwise might not be able to detect the crossing in advance. They may also have trouble detecting an approaching train because of sight obstructions in the vehicle or because of uncertainty in determining where to look along the tracks.

Crossing-angle categories of 0 to 29 degrees, 30 to 59 degrees, and 60 to 90 degrees were analyzed to determine their influence on the effectiveness of warning systems. Data on the influence of angle of crossing on the effectiveness of warning systems are given in Table 4 and shown in Figure 3. Review of the data in Table 4 indicates that for two upgrade categories, passive to fixed distance and passive to constant warning time, effectiveness values in the

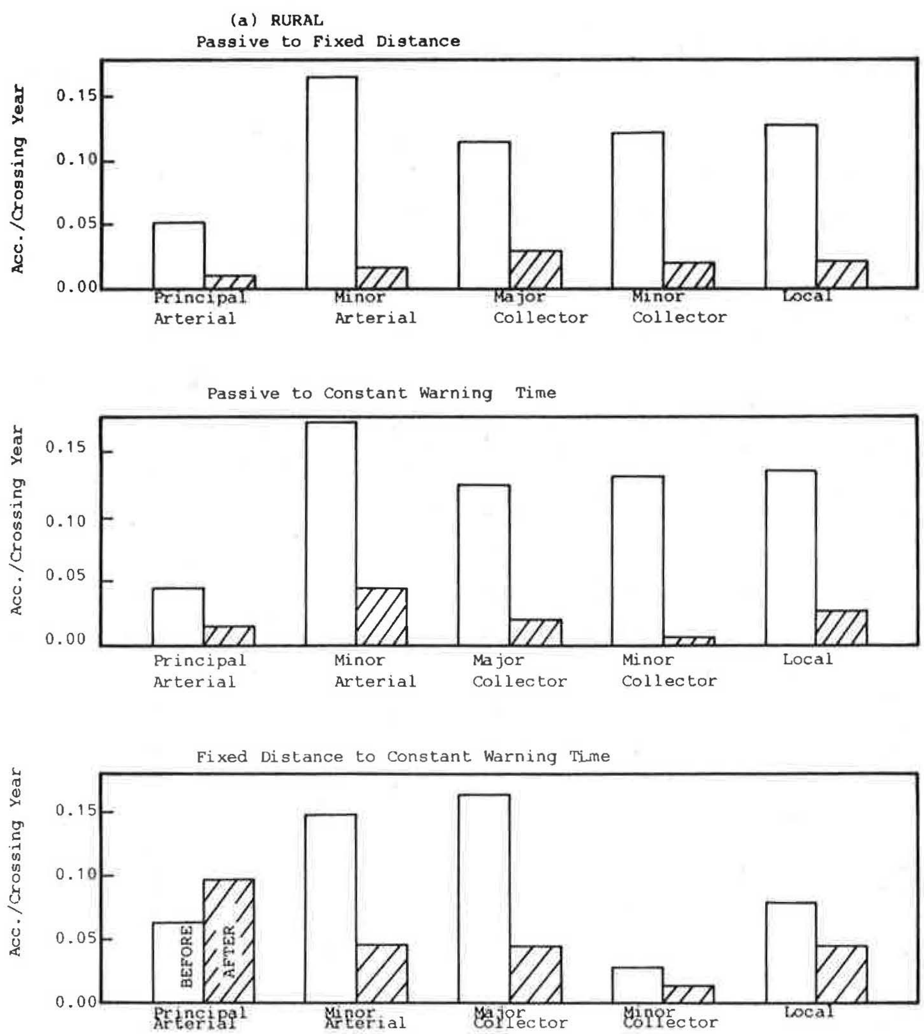


FIGURE 2 Graphical comparison of before-and-after accident rates for active system upgrades as a function of road classification.

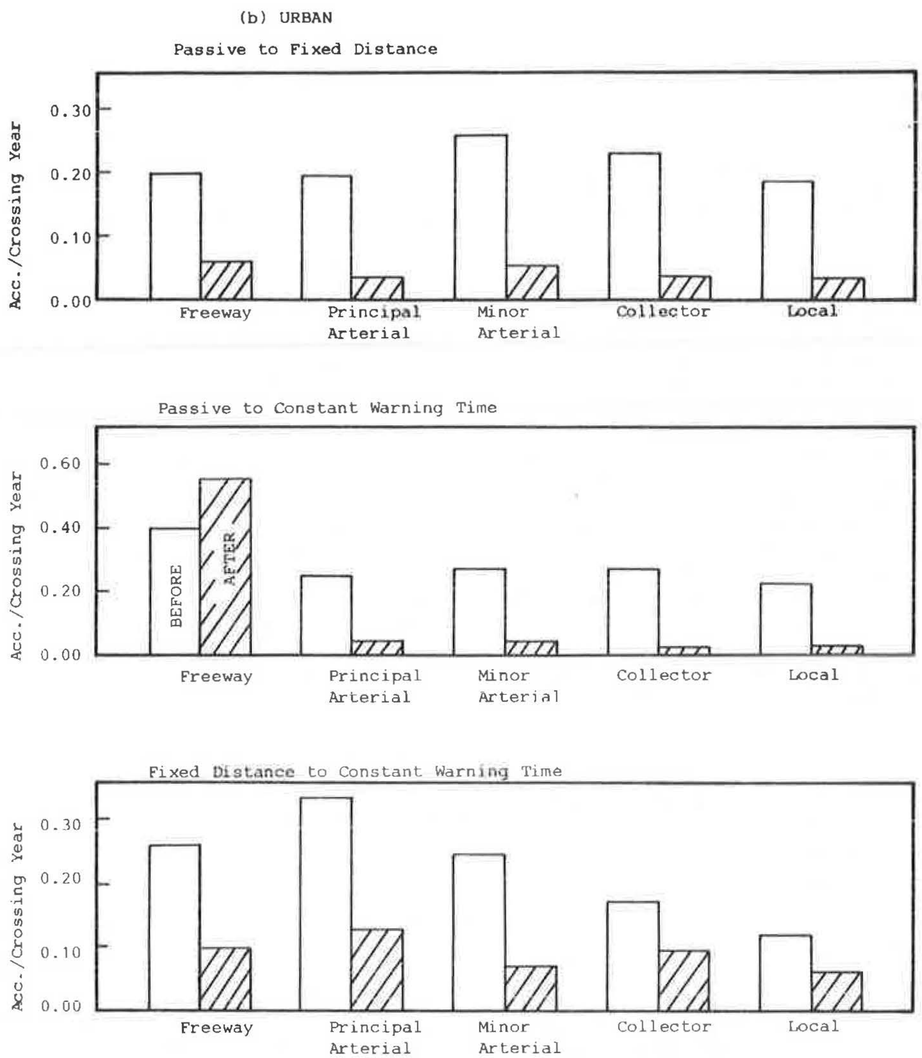


Figure 2 continued.

60-to-90-degree-angle category are roughly equal to or greater than those in the oblique-angle categories. This is contrary to the hypothesis stated at the beginning of this section. As expected, the effectiveness of upgrades in the fixed-distance-to-constant-warning-time category was greatest in the 0-to-29-degree-angle class, in which there was an effectiveness value of 68 percent. In most cases the confidence intervals were rather large because of the relatively small sample sizes. There was an overlap in the confidence intervals of the three upgrade angle categories; this indicated that there was no significant difference between effectiveness values at the 95-percent confidence level.

Although the lack of any definite pattern as far as variation in effectiveness with angle of crossing was unexpected, the results may be explained by two items of information not included in the data base. The first of these is sight distance. Sight distance is not quantified in the data base and sight obstructions (such as vegetation or structures) are not noted. The second factor is the direction of ap-

TABLE 4 Influence of Angle of Crossing on Effectiveness of Active Warning Systems

Upgrade Category	Angle of Crossing (degree)	No. of Crossings	Before Upgrade Accidents	Before Upgrade Crossing Years	After Upgrade Accidents	After Upgrade Crossing Years	Effectiveness Value (%)	95 Percent Confidence Interval (%)
Passive to fixed distance	0–29	189	148	695	29	814	83	77 to 89
	30–59	330	163	1,200	41	1,448	79	72 to 86
	60–90	1,639	940	6,120	188	6,942	82	79 to 85
Passive to constant warning time	0–29	81	57	263	15	379	82	72 to 92
	30–59	75	33	229	9	358	83	70 to 96
	60–90	441	256	1,454	51	2,032	86	82 to 90
Fixed distance to constant warning time	0–29	37	19	75	11	134	68	46 to 90
	30–59	54	32	175	29	333	52	30 to 74
	60–90	321	128	763	103	1,510	59	49 to 69

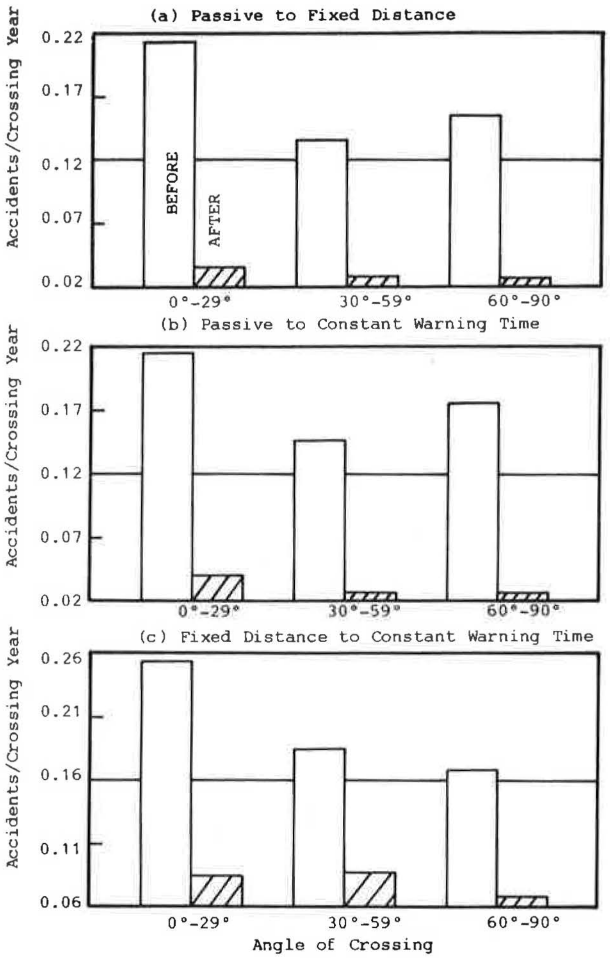


FIGURE 3 Graphical comparison of before-and-after accident rates for active system upgrades as a function of crossing angle.

proach of vehicular and train traffic. Both of these factors are important in determining when drivers first detect the presence of trains, yet neither is included in the data base.

Influence of Train Speed

The introduction to this paper alluded to the motorist credibility problem associated with fixed-distance systems at crossings at which there is significant variation in train speed. At locations where the time interval between signal activation and actual passage of the train through the crossing is lengthy, some drivers may try to proceed past flashing lights or maneuver around crossing gates. Because they provide a shorter and more uniform waiting time, constant-warning-time systems would be expected to be more effective than fixed-distance systems at crossings at which there are large variations in train speed.

It should be pointed out that warning device credibility is a function of the track circuit design speed and the frequency distribution of actual train speeds. Neither variable was considered in this study. The inventory data base did not include information on track circuit design speed. A preliminary analysis was made of the actual train speeds reported in the accident data base. It became apparent that, in some cases, the reported speed was not accurate, and many times speed data were not reported at all.

The influence of train speed on the effectiveness of warning systems was examined by using three different measures. One concept was the speed-difference approach, in which speed difference was the algebraic difference between maximum timetable speed and typical minimum speed. Second, a speed-ratio concept was examined, in which the ratio of maximum timetable speed to typical minimum speed was computed. It was hypothesized that large speed differences and large speed ratios would be associated with high accident rates (or greatest danger) for crossings equipped with fixed-distance warning systems. Finally, a maximum timetable speed was used to check the effectiveness of the two different warning systems, because high speed is usually associated with the highest accident rates.

Figure 4 shows the relationship between train speed difference and effectiveness values for the three upgrade categories being considered. It was recognized that these points do not represent continuous data; nevertheless, this format was chosen to make it easier to identify any trends that might exist. Although the fixed-distance-to-constant-warning-time upgrade had the lowest effectiveness value, it was the only upgrade category to show an increasing trend when tested using Spearman's Rho statistical test. This tends to confirm the hypothesis stated earlier in this section; that is, the effectiveness of constant-warning-time systems should increase as variation in train speed increases. To provide additional insight, graphical comparisons of the before-and-after accident rates, on which the effectiveness values were calculated, were made (see Figure 5).

Figures 6 and 7 present the relationship between speed ratio and effectiveness values for the three upgrade categories. Spearman's Rho tests revealed no significant trends in any of the upgrade categories. To investigate the relationship further, train speed ratios were grouped differently than those shown in Figure 6 (results are not shown here). There was still no significant relationship between train speed ratio and effectiveness values.

The relationship between maximum timetable speed and effectiveness value is shown in Figures 8 and 9. Spearman's Rho revealed no significant trends in any of the upgrade categories. Thus it was concluded that neither speed ratio nor maximum timetable speed had an influence on the effectiveness of warning systems.

CONCLUSIONS

This study carried the analysis of the rail-highway grade-crossing inventory and accident data bases one step farther than had been done in the past. Rather than considering the effectiveness of active warning devices in general, effectiveness values were developed for fixed-distance and constant-warning-time systems for several different stratifications of variables. Upgrades from passive devices to fixed-distance and constant-warning-time systems had almost equal effectiveness values--82 and 85 percent, respectively. For changes from fixed-distance to constant-warning-time systems, the effectiveness value was 26 percent. This result tended to confirm the hypothesis that constant-warning-time systems have greater credibility with motorists than do fixed-distance systems.

Other important conclusions drawn from the study are as follows:

1. Functional class of road had no apparent influence on the effectiveness of warning systems

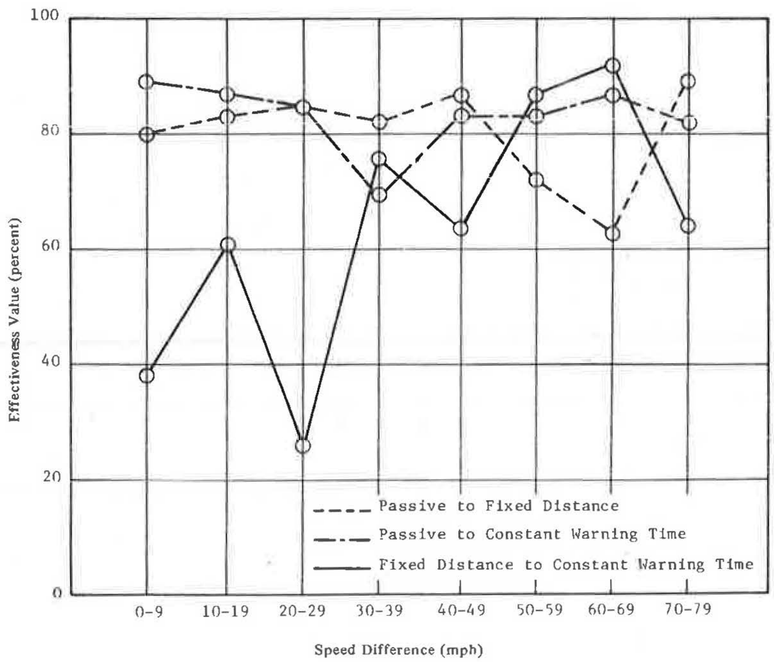


FIGURE 4 Relationship between effectiveness values and train speed difference for constant-warning-time and fixed-distance system upgrades.

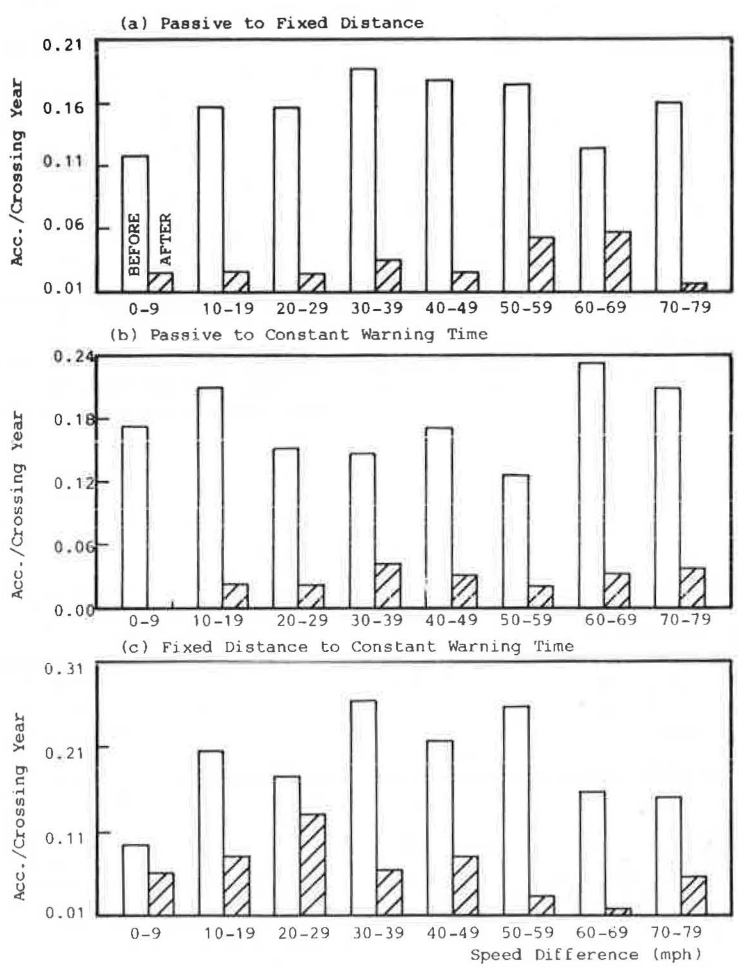


FIGURE 5 Graphical comparison of before-and-after accident rates for constant-warning-time and fixed-distance system upgrades as a function of train speed difference.

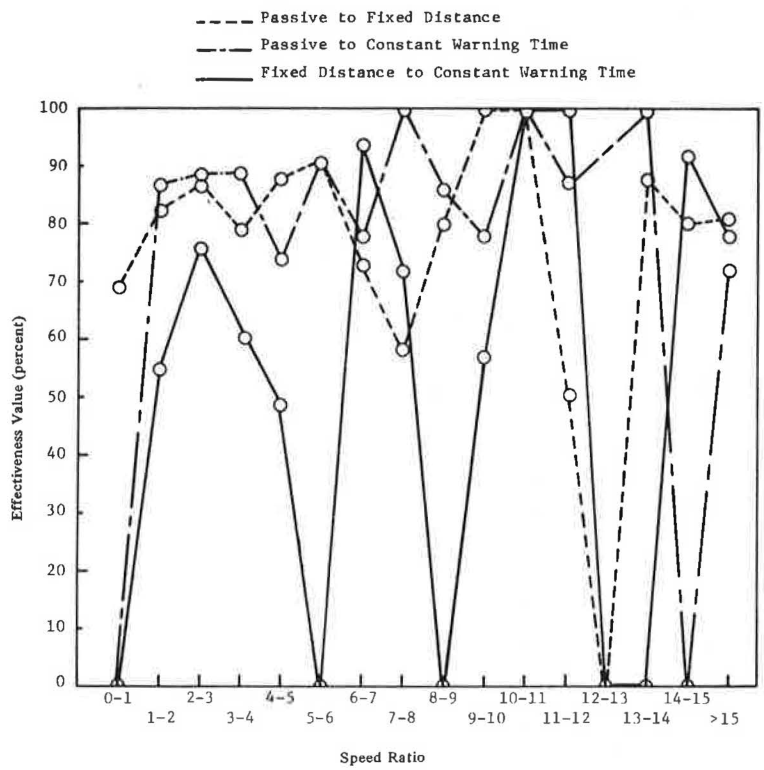
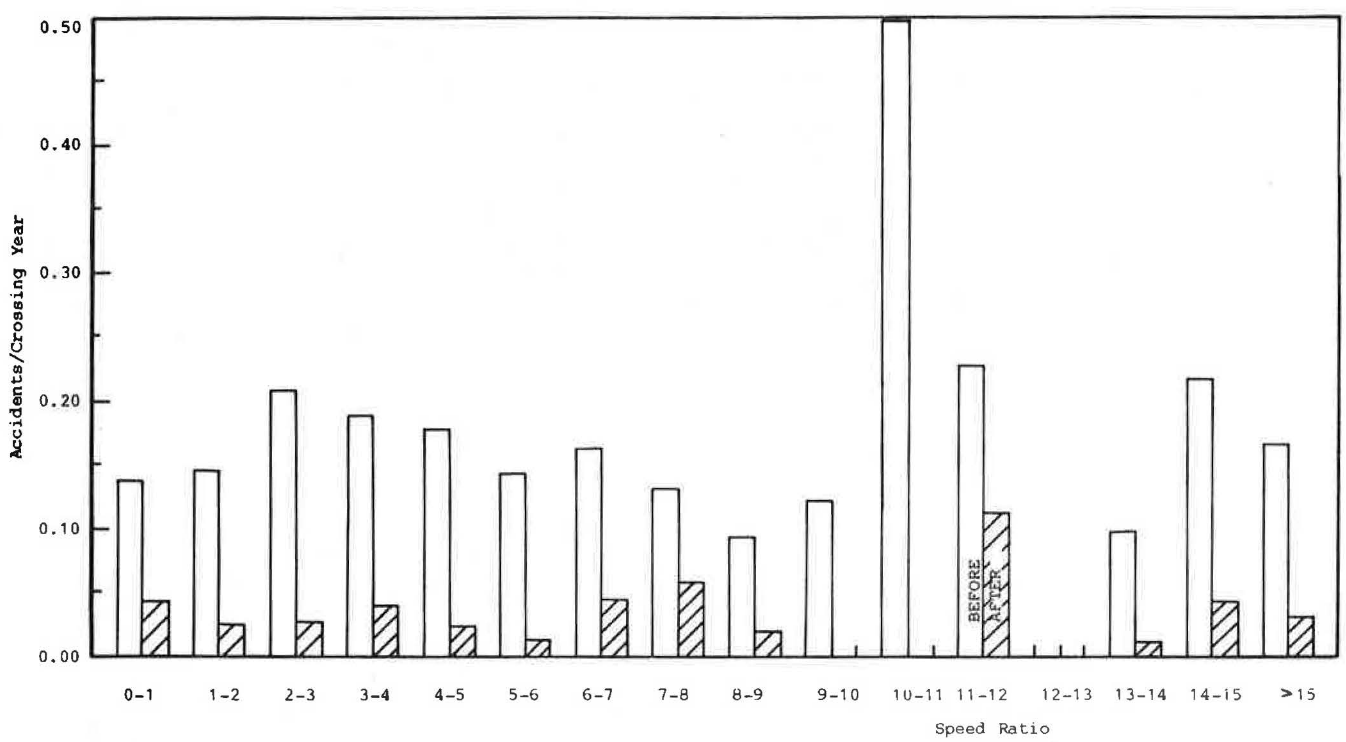
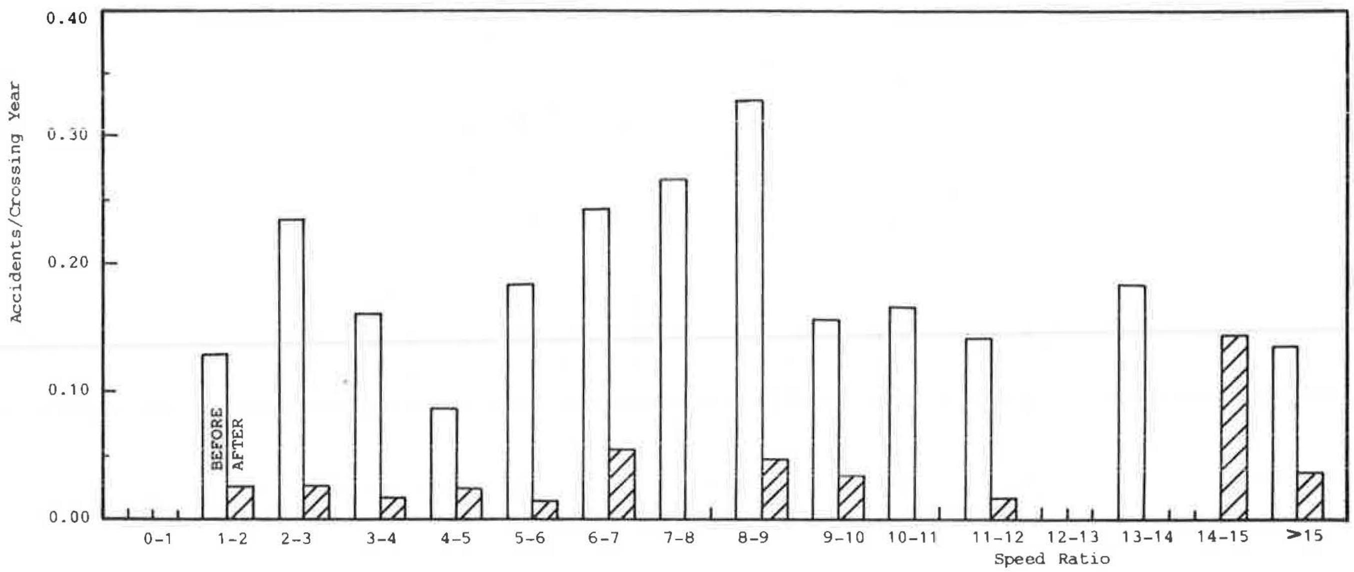


FIGURE 6 Relationship between effectiveness values and train speed ratio for constant-warning-time and fixed-distance system upgrades.

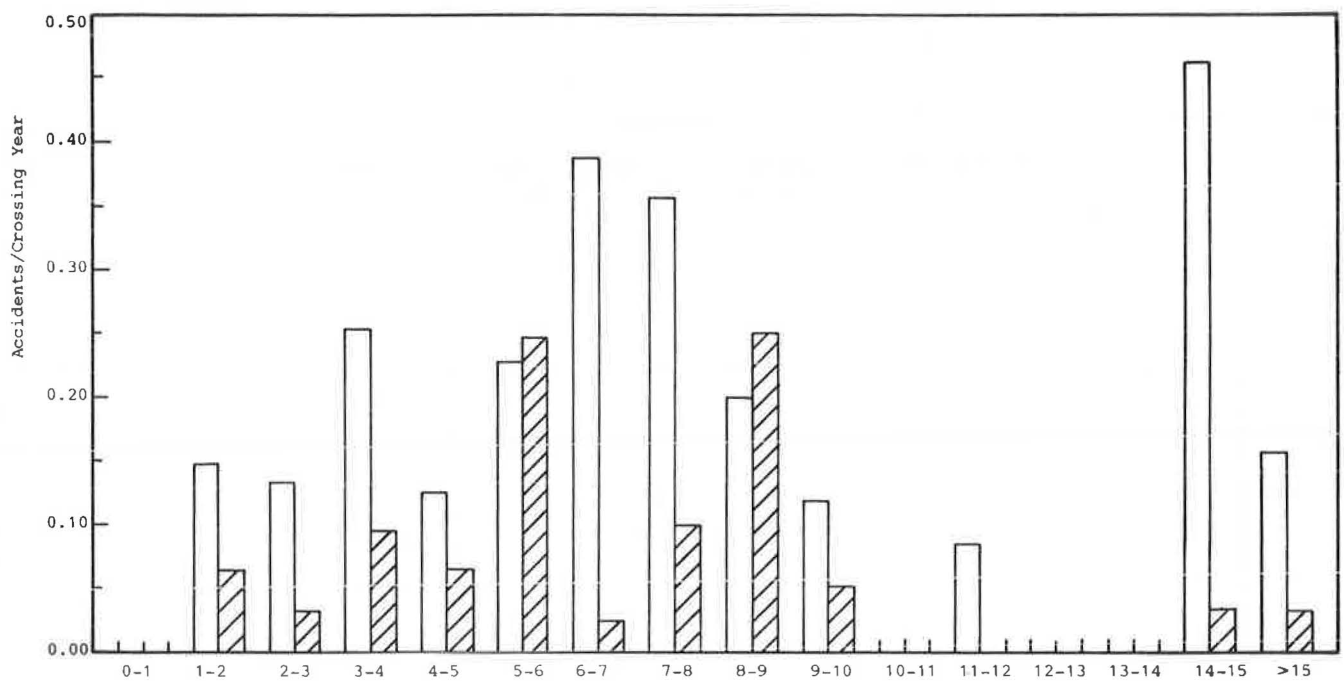


(a) Passive to Fixed Distance

FIGURE 7 Graphical comparison of before-and-after accident rates for constant-warning-time and fixed-distance system upgrades as a function of train speed ratio.



(b) Passive to Constant Warning Time



(c) Fixed Distance to Constant Warning Time

Figure 7 continued.

for fixed-distance and constant-warning-time upgrades.

2. For passive-to-fixed-distance and passive-to-constant-warning-time upgrades, effectiveness values in the 60-to-90-degree-angle category were essentially equal to or slightly greater than those in the oblique-angle categories (approximately 82-percent effectiveness).

3. As expected, the effectiveness of upgrades in the fixed-distance-to-constant-warning-time category was greatest for the angle-of-crossing class of 0 to 29 degrees, which had 68-percent effectiveness.

4. A significant relationship was found between train speed difference and constant-warning-time systems; that is, system effectiveness increased as the variation in train speeds at a location increased.

5. Train speed, as measured by the speed ratio and maximum timetable speed, had no apparent influence on the effectiveness of warning systems for fixed-distance and constant-warning-time upgrades.

RECOMMENDATIONS

Results of this study suggest several areas in which additional research is recommended. These are outlined in the following paragraphs.

Additional research is warranted to analyze the DOT-AAR data files to determine if, when normalized

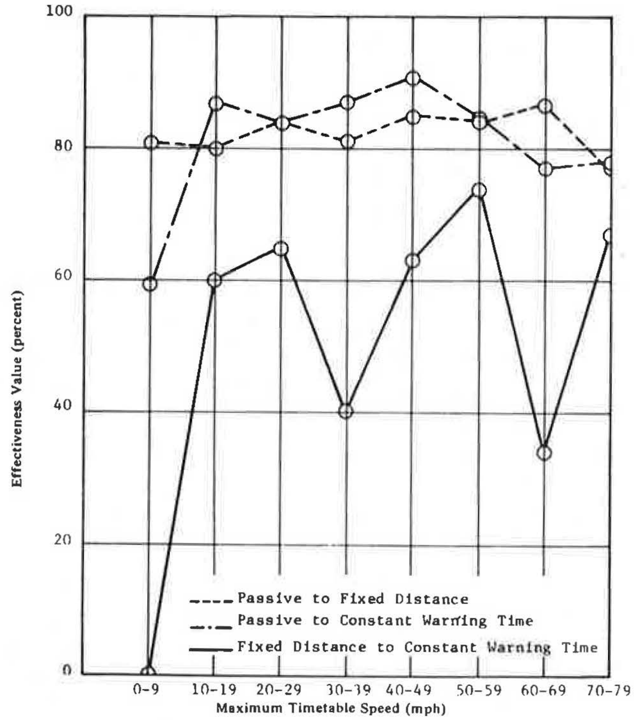


FIGURE 8 Relationship between effectiveness values and train maximum timetable speed for constant-warning-time and fixed-distance system upgrades.

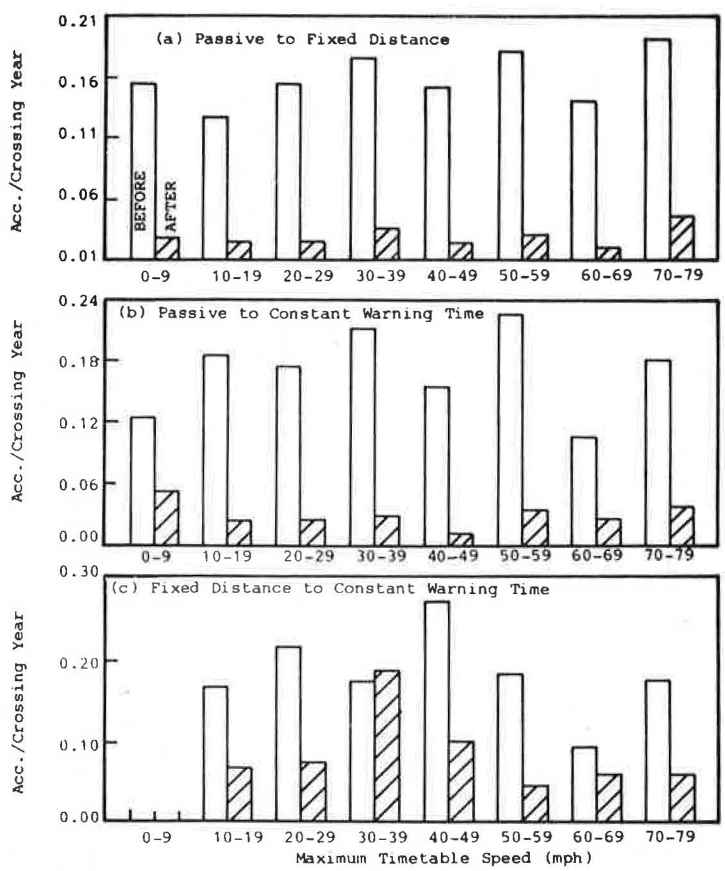


FIGURE 9 Graphical comparison of before-and-after accident rates for constant-warning-time and fixed-distance system upgrades as a function of maximum timetable speed.

by exposure, accident rates at crossings with fixed-distance systems differ from those at crossings with constant-warning-time systems. Only by normalizing the accident rates by exposure (traffic volume times train volume) can the credibility issue be addressed. In this way, the hypothesis that accident rates at crossings equipped with fixed-distance systems would be expected to increase with increasing difference between typical train speeds and maximum timetable speeds could be tested.

Similarly, it would be desirable to make the comparisons described in this paper for different exposure levels. The authors are currently conducting analyses of accident rates by exposure. Although the results are not yet in a form suitable for inclusion in this paper, it is anticipated that they will be published at a later date.

Another area of future research would be the development of statistical models to identify variables that are significantly related to grade-crossing accident rates (normalized by exposure) for fixed-distance and constant-warning-time systems. Identification of such factors would be useful in refining guidelines or warrants for installing fixed-distance and constant-warning-time systems. In addition, the development of capital and life-cycle costs of the two different warning systems would provide another source of input for the development of installation guidelines.

ACKNOWLEDGMENTS

This study relied heavily on data contained in the national DOT-AAR inventory and FRA accident files. The assistance of the West Virginia Department of Highways (WVDOH) in cooperation with the FHWA, U.S. Department of Transportation, in sponsoring a previous study that established the data base at West Virginia University is acknowledged. Special appreciation is expressed to Ray Lewis (WVDOH) and Janet Coleman (FHWA) for their assistance and encouragement throughout the study.

Discussion

Brian L. Bowman*

Halkias and Eck are to be commended for identifying the need for, and their willingness to conduct, an independent study on the effectiveness of constant-warning-time devices. Determinations on the effectiveness of improvements at rail-highway grade crossings is a difficult undertaking. The task is made complex by the relatively low number of accidents that involve trains, the accuracy of requisite operational and physical data, and the determination of appropriate exposure factors.

Review of the study effort prompts the following comments:

1. The DOT-AAR Crossing Inventory File was used to provide information on crossings for the study. This file provides the only means to obtain national information on the physical and operational characteristics of crossing without contacting individual railroads and states. The inventory requires the active support of both railroads and states to maintain current data. Often this support is not as universal as desired and changes take place without accompanying inventory updates. The result is that the inventory, although probably the best tool available for obtaining information on crossings, does not always contain current and accurate data.

This problem has been found to be prevalent in train speeds, traffic and train volumes, and the entry used to designate the presence of train speed selection equipment (i.e., constant-warning-time devices). Some of these discrepancies become evident when the inventory is closely scrutinized. For example, the inventory was searched for all crossings that had (a) a positive response to speed selection capabilities, and (b) only passive warning devices. This search revealed that more than 201 crossings have constant-time-control capabilities in conjunction with passive warning devices. This result is contradictory because if train detection equipment exists there are probably active warning devices present at the crossing.

The authors have recognized these difficulties and provided partial control by including only crossings with positive responses to both speed selection and active warning. This will serve to eliminate the erroneous passive warning entries. However, those crossings with active warning devices and erroneously coded as having speed selection capabilities are still included in the study. Unfortunately, the only solution to this problem is to verify that the correct combinations of detection and warning devices exist on a site-by-site basis. This would be a huge task and, possibly, outside the scope of the authors' study. Without the verification, however, it is unknown if we are actually analyzing crossings with the desired combination of detection and warning devices.

2. As mentioned by the authors, constant-warning-time devices are intended to prevent train accidents that are attributable to driver impatience. Therefore, these accidents would be characterized by vehicles being impacted by or striking the first unit of the train. For example, the installation of constant-warning-time devices would not be expected to reduce the number of accidents in which the tenth consist of the train is impacted. This type of accident indicates that (a) the vehicle was not stopped at the crossing, (b) the driver was not subjected to an excessive wait time, and (c) driver impatience was not a factor in the accident.

The measures of effectiveness chosen for an evaluation should have at least a casual relationship with the project objectives. Because the study analyzed total number of accidents without consideration to specific accident types, there is an uncertainty as to the proper interpretation of the study results.

3. The authors performed comparisons between analysis groups without investigating the need to stratify sites by physical and operational characteristics. Consideration should have been given as to why constant-warning-time devices are installed to determine if stratification of analysis sites is required. If, for example, the devices are primarily installed to alleviate problems caused by large train speed ratios, then all analysis categories should possess the same train speed ratio. The failure to stratify creates no problems as long as comparisons within groups, such as before-and-after analysis on the same sites, are performed. If, however, analysis between groups that have different

*Goodell-Grivas, Inc., 17320 West Eight Mile Road, Southfield, Mich. 48075.

physical and operational characteristics (i.e., fixed distance with flashing lights versus constant warning with flashing lights) takes place without stratification, then the results can be confounded. Thus the conclusions of this study, which are based on comparisons between groups without investigation of the need for stratification of analysis sites, should be interpreted with caution.

4. The authors used confidence intervals to infer significance by inspecting the data range and to compare data from different populations. The use of confidence intervals is good and actually provides more information than only reporting a hypothesis test or a significance level. However, statistical significance should not be inferred by inspecting the range that exists between the confidence limits. This range is established by relationships between data items within the analyzed sample and the sample size. Observing values outside of the confidence band indicates a relatively unlikely event, given the hypothetical situation analyzed. Stating that significance exists because zero is not within the limit is misleading. Similarly, comparisons of confidence limits between different analysis groups, with different physical and operational characteristics, are confounded and also misleading.

5. The authors hypothesized that the effectiveness of constant-warning-time devices would be expected to be related to crossing angle. Whether this would be expected or not is questionable and the question is not resolved by the results of this study. Changes from fixed-distance systems to constant-warning-time systems were analyzed by grouping sites with flashing lights and sites with gates together. The presence or absence of gates may, however, have a greater influence on accidents than the type of train detection circuitry. Failure to identify the degree of improvement that was attributable to constant-warning-time devices and that which was due to gate installation precludes any conclusions on the effectiveness of constant-warning-time upgrades with respect to different crossing angles.

6. The need for using exposure factors, considering both roadway and train volumes, was identified by the authors. This is especially important because the analysis consisted of accidents occurring during a 7-year period. A considerable amount of change, both in roadway and train volumes, can be expected to occur in a 7-year period. This change should be accounted for by using exposure factors or controlled by employing either comparative or control site experimental plans. Because this was not done in the study, it is not known what portion of the observed change is caused by the analysis variables and what portion is caused by changes in train and traffic volumes.

In summary, the authors have identified an issue that is in need of further research. Constant-warning-time devices are often installed because "everything else has been tried." Knowledge about their effectiveness will enable device deployment based on their probable effect and not on intuitive judgment. This, however, requires a strong experimental design to minimize validity threats and confounding effects. The need for a stronger evaluation has been identified by the authors, but the resources for such an evaluation were probably beyond the scope of this study. The applicability of the conclusions and effectiveness factors are therefore constrained, and caution should be exercised in interpreting the study results.

Discussion

William D. Berg*

The research conducted by Halkias and Eck has revealed that the use of constant-warning-time track circuits can have a positive influence on safety. However, the level of effectiveness that can be expected under various real-world conditions remains undetermined. Several comments will be offered about interpretation of the findings presented by Halkias and Eck, as well as about the direction of future research in this area.

Data on the overall effectiveness of constant-warning-time track circuits are presented in Table 2 and Figure 1. Because the estimated effectiveness for upgrades to flashing lights or gates with a constant-warning-time track circuit fall within the 95-percent confidence interval for the corresponding upgrades with a fixed-distance track circuit, the effectiveness of the track circuit design cannot be distinguished from the obvious benefits created by the upgrade to an active warning device. For those crossings at which only a change in track circuit occurred, the confidence interval for the effectiveness factor includes zero for both flashing-light and gate systems. Thus, based on the data presented by the authors, it cannot be concluded that type of track circuit has a statistically significant impact on safety. This does not mean that no benefit exists, but simply that the data base was not able to permit its measurement. When the data are aggregated over both flashing-light and gate systems, the estimated 26-percent effectiveness of constant-warning-time track circuits is significant. This supports the hypothesis that this type of track circuit can be more effective than traditional fixed-distance designs, but it does not explain under what conditions these benefits can be expected to occur.

Examination of Figures 1a and b reveals that the after-accident rates for both flashing-light and gate upgrades are approximately equivalent, regardless of track circuit design. In addition, the before-accident rates for gate upgrades are larger than for the flashing-light upgrades. This suggests that the effectiveness of automatic warning devices, as measured by the actual change rather than the percentage change in accident rate, is principally a function of the before-accident rate. Restated, automatic warning devices appear to provide a given absolute level of safety, and the relative change to that level depends on the accident rate that existed before the improvement.

The comparison of the accident rates shown in Figure 1c indicates that constant-warning-time track circuits can provide an accident reduction of about 0.03 to 0.04 accident per year (although as noted previously these estimates are not statistically significant). In addition, for those flashing-light crossings that received an upgrade to a constant-warning-time track circuit, the data suggest that a substantially greater accident reduction could have been achieved if gates had been installed and no change had been made to the track circuit. This would certainly be intuitively reasonable. Finally, the grade crossings represented in Figure 1c exhibit much larger accident rates (both before and after the change in track circuit) than the similarly

*Department of Civil and Environmental Engineering, University of Wisconsin-Madison, 1415 Johnson Drive, Madison, Wis. 53706.

equipped crossings represented in Figures 1a and b. This probably reflects substantially different exposure levels, because the former crossings were upgraded to active warning devices at an earlier time because of high train and traffic volumes. This observation also points out the importance of incorporating exposure in accident rate calculations and comparisons.

The data presented in Figure 3 show that grade crossings with acute angles of less than 30 degrees are more hazardous that those with larger acute angles. As noted by the authors, the principal influencing factor is probably the corner sight distance at the crossing, a factor that is not available in the data base. The implied 52- to 68-percent effectiveness for track circuit upgrades is misleading and should be considered unreliable because, of the 412 grade crossings in the sample, 54 percent involved a concurrent upgrade from flashing lights to gate. The effectiveness of this improvement virtually obscures the benefits of the more responsive track circuit. As noted previously, Figure 1c suggests that constant-warning-time track circuits can reduce accident rates by 0.03 to 0.04 accident per year.

The data in Figures 4-9 reveal that speed difference and speed ratio do not provide any useful insight into the effectiveness of constant-warning-time track circuits. This should not be unexpected because the benefits of these train detection systems is due to their credibility, and this is a function of the track circuit design speed and the range of actual train operating speeds. The speed-difference and speed-ratio variables are poor indicators because they rely on maximum train speed rather than on track circuit design speed. The typical credibility problem occurs when the maximum train speed over a crossing is reduced without a concurrent change in the track circuit. This causes an increase in warning times beyond the desired 25-sec time interval, thereby creating a situation in which there sometimes is more than ample time for a motorist to safely traverse the crossing even though the warning devices are operating. The greater the difference between the track circuit design speed and the minimum train speed, the greater the credibility problem. A constant-warning-time track circuit virtually eliminates the credibility problem.

In conclusion, the research conducted by Halkias and Eck does tend to confirm the hypothesis that constant-warning-time track circuits can provide greater safety when conditions warrant their use. However, it is doubtful that estimates of the magnitude of these benefits are necessary for resource allocation studies for two reasons. First, existing accident prediction procedures are not sufficiently accurate to reliably distinguish the small differences in accident rates associated with alternative track circuit designs. Second, decisions regarding the type of track circuit that should be used at a crossing are quite properly a design decision rather than a resource allocation decision. Guidelines for selecting type of track circuit, as well as placement of signals (cantilevered versus mast mounted), are already available (7). Therefore, it is not clear that further research on the effectiveness of constant-warning-time track circuits will lead to useful and implementable results. If additional work is to be conducted, it should be based on data that include the design speed of fixed-distance track circuits and accident rates normalized for exposure, and the experimental design should use a treatment-control type of before-and-after comparison (8).

Discussion

John S. Hitz*

Thank are extended to the Committee on Railroad-Highway Grade Crossings for this opportunity to comment on the paper by Halkias and Eck. This subject is of great personal interest to me because I have been involved in similar research at the Transportation Systems Center (TSC), U.S. DOT, during the past several years. Comments are addressed in particular to determining the effectiveness of constant-warning-time devices.

Efforts to determine the effectiveness of constant-warning-time devices are worthwhile. Because these devices add significantly to the costs of warning device improvement projects, their use should be justified by a resultant increase in effectiveness. If they can be shown to be cost-effective, then they constitute an additional "weapon" in the arsenal of preferred means of improving crossing safety. It should be mentioned that these devices have additional benefits in their ability to improve highway traffic flow, which further justifies their use in certain applications.

The Halkias and Eck study determined an average effectiveness of 26 percent for constant-warning-time-device additions to flashing lights and gates compared with fixed-distance systems. The 95-percent confidence interval for this value is quite large, however. The true value of effectiveness could lie anywhere between 3 and 49 percent. This large uncertainty is a reflection of the small amount of data available for analysis. However, practical insight of crossing safety suggests that these devices should have some positive level of safety improvement. Increasing the credibility of warning devices should result in fewer instances of motorists taking risks to avoid long waits at railroad grade crossings. Results of a similar study at TSC tend to support this notion and are consistent with the Halkias and Eck study. At TSC it was found that the effectiveness of all flashing lights and gates tended to be lower at crossings with large variations in train speed. Although the results of these studies suggest that constant-warning-time devices are effective, it would be desirable to have more confident answers on this issue. I would like to provide some suggestions on how it is possible to move toward this goal through further analysis of the available data. Any such study, however, must recognize and resolve to the extent possible problems with both the quality and quantity of the data.

The limited quantity of data available on constant-warning-time devices lowers the confidence that can be placed in resulting effectiveness values. This problem is aggravated when the data are sectionalized to analyze specific factors that influence effectiveness, as is inevitably the case. Therefore, as much as possible of the data that are available for analysis should be used. This can be accomplished by concentrating further analysis on the data for upgrades from passive devices to lights and gates that do and do not include constant-

*Transportation Systems Center, DTS-54, U.S. Department of Transportation, Kendall Square, Cambridge, Mass. 02142.

warning-time devices. This is the largest group of upgrades and will thus yield the most confident results. Halkias and Eck investigated this group but found no difference in effectiveness between upgrades that did and did not include constant-warning-time systems. These results should be investigated further by addressing some of the following issues regarding data quality.

Several problems with data quality result from an inability of the data to fully describe features of crossings that may influence the effectiveness of warning devices (e.g., restricted sight distance). If constant-warning-time devices are systematically chosen for installation at crossings with restricted sight distance, the data may reveal the devices to have a lower-than-actual level of effectiveness. This problem can be minimized by ensuring that the two groups of crossings being compared (upgrades to fixed-distance and constant-warning-time devices) are equivalent in terms of potential for accidents before the upgrade. This will tend to control for those factors not in the inventory that may influence the hazard level of a crossing and thus the effectiveness of warning devices. It is recommended that the DOT basic accident prediction formula be used because it is the best indicator of the hazard level of the crossing before upgrade. This does not necessarily require that the crossings in each group be categorized into subgroups of equal hazard, which would reduce sample sizes. A reasonable requirement would simply be that the two groups have the same distribution of hazard levels.

A similar data quality problem is failure of the data to describe the full extent of improvements that may take place when a constant-warning-time device is installed. For example, flashing lights may frequently be replaced with larger, more effective lights at the same time that constant-warning-time devices are installed. With the data available it is difficult to determine if resultant safety improvements are caused by the improved lights or by the addition of constant-warning-time devices. This problem will be largely avoided by investigating upgrades from passive devices because only new lights of similar effectiveness will be involved.

Another problem with data quality to be addressed is the vagueness with which constant-warning-time devices are defined. The existence of a constant-warning-time device can only be implied from the data by a positive response to the ambiguous question, Does crossing signal provide speed selection for trains? The type of constant-warning-time device is not indicated. Many of the devices could be of the motion-detector type. These devices are intended more to reduce long traffic delays and congestion than to provide a constant warning time. Therefore, they would tend to have little impact on reducing accident statistics, because the crossings involved would generally have low-speed switching movements and few accidents to begin with. If significant numbers of motion detectors are included in the data, then the effectiveness results for constant-warning-time-devices could be biased downward. To reduce the occurrence of this problem, the constant-warning-time-device group should be screened to eliminate most motion detectors by excluding crossings with large numbers of switch trains or primarily low-speed trains or both.

Regarding train speeds, a more detailed analysis of train speed variation should be enlightening. Do crossings in the constant-warning-time group actually have large variations in train speed? How does this compare with the same information for the fixed-distance group? One would expect that the constant-warning-time group would have the greatest train speed variation. When determining the effectiveness of a constant-warning-time device relative to a fixed-distance device, the comparison should be made between locations with equivalent levels of train speed variation. The basic question that is being addressed is whether constant-warning-time devices are more effective than fixed-distance devices under conditions of large train speed variations. If level of train speed variation is not controlled, the results could be significantly biased.

Another type of data limitation results from possible changes to crossing characteristics that may influence effectiveness after a warning device upgrade has taken place. For example, significant changes to train and highway traffic could occur after an upgrade, thus increasing the likelihood of an accident. Anticipated changes in such characteristics of crossings may lead to some decisions about upgrades. If the inventory does not account for these changes or if they are not considered in the analysis, effectiveness results could be biased. Unfortunately, this is a particularly difficult problem to overcome. Even if various editions of the inventory are analyzed to determine crossing changes over time, there is no assurance that the actual changes that have taken place have been reported.

If the precautions that are outlined in this discussion are considered, it is possible that the results may have a sufficiently high level of confidence to be of practical use. In any event, the data will have been used for all its useful information. The suggestions in this discussion are consistent with proposals by Halkias and Eck for future work. I wish them success and look forward to working with them in these efforts.

Authors' Closure

We greatly appreciate the thoughtful and constructive reviews made by Hitz, Bowman, and Berg of our paper. We agree with their comments concerning clarification of certain items in the national data base and on the need for a sound experimental design (including choice of appropriate variables) in any work of this nature. Although we were remiss in neglecting certain critical points in our analysis (for example, analysis of total accidents without consideration of specific accident types) lack of resources constrained us in other areas, most notably the site-by-site verification of the existence of the correct combination of detection and warning devices.

The only specific issue we wish to address concerns Bowman's comments on statistical significance and comparisons of confidence limits. Stating that significance exists because zero is not within the confidence limit is valid, as in the case of the 26-percent effectiveness of fixed-distance to constant-warning-time upgrades at a specified confidence level. For situations such as the one encountered here, that is, working with ratios, comparisons of confidence limits between two sample means are appropriate and valid tests, because confidence intervals not only provide information on the true mean of a sample but are also used as hypothesis tests for differences between means.

We recognize that, in general, our paper may have raised more questions than it answered. Although the effectiveness factors developed may not be directly applicable at this point in time, the process of producing these "first-cut" effectiveness factors has resulted in new information regarding the use and effect of different types of grade-crossing warning devices. We did not mean to imply in our recommendations that decisions regarding what type of track circuit to use were resource allocation decisions. As Berg correctly pointed out, they are design decisions. The point to be made is that even though the designer may not use the actual quantities developed in our work, the insights provided by our findings should lead to improved decision making.

Perhaps the greatest value of the paper, and the discussion that has taken place relative to it, is the explicit identification of data limitations and areas in which additional research is needed. It is hoped that the identification of these limitations will encourage maintenance of a current and accurate DOT-AAR Crossing Inventory File and even serve as an impetus to making some minor modifications to the data base that will enhance its use as a decision-making tool. Identification of research needs should prove of interest to both researchers and funding agencies.

Based on the information in this paper and in the discussions, the major data base and research issues relative to rail-highway grade crossings in general and fixed-distance versus constant-warning-time systems in particular have been outlined. These are as follows:

1. The inventory file does not always contain current and accurate information.
2. Certain additional crossing features (mainly sight distance and extent of improvements made) should be added to the data base.
3. The definition of constant-warning-time devices in the data base needs to be improved.
4. It is imperative that exposure be incorporated into accident rate calculations and comparisons.
5. The use of track circuit design speed rather than maximum train speed is needed to provide insight into the effectiveness of constant-warning-time systems.
6. A treatment-control type of before-and-after experimental design is recommended for work of this nature.
7. The comparisons described in this paper should be made for different stratifications of physical and operational characteristics (including exposure).
8. In examining device effectiveness versus angle of crossing, the degree of improvement attributable to constant-warning-time systems and that due to gate installation must be determined.

REFERENCES

1. J.A. Coleman and B.F. George. National Railroad-Highway Crossing Inventory. Public Roads, Vol. 47, No. 2, Sept. 1983, pp. 66-68.
2. E. Farr and B.H. Tustin. Optimizing Resources at Rail-Highway Crossings. ITE Journal, Vol. 52, No. 1, Jan. 1982, pp. 25-28.
3. E.H. Farr and J.S. Hitz. Effectiveness of Motorist Warning Devices at Rail-Highway Crossings. Draft Report. Transportation Systems Center, U.S. Department of Transportation, Cambridge, Mass., Sept. 1983, 84 pp.
4. J. Morrissey. The Effectiveness of Flashing Lights and Flashing Lights and Gates in Reducing Accident Frequency at Public Rail-Highway Crossings, 1975-78. Report FRA-RRS-80-005. FRA, U.S. Department of Transportation, Jan. 1981, 20 pp.
5. E.H. Farr and J. Hitz. Additional Investigation into Rail-Highway Crossing Warning Device Effectiveness. FHWA, U.S. Department of Transportation, April 1982, 32 pp.
6. R.W. Eck and J.A. Halkias. Effectiveness of Warning Devices at Rail-Highway Grade Crossings. Department of Civil Engineering, West Virginia University, Morgantown, May 1984, 69 pp.
7. Rail-Highway Grade Crossing Warning Systems and Surfaces. The Railway Progress Institute, Alexandria, Va., 1983.
8. W.D. Berg. Experimental Design for Evaluating the Safety Benefits of Railroad Advance Warning Signs. Report FHWA-RD-79-78. FHWA, U.S. Department of Transportation, April 1979.

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

Procedure for a Priority Ranking System for Rail-Highway Grade Crossings

TIMOTHY A. RYAN and JOHN W. ERDMAN

ABSTRACT

In 1982 the Interstate Division for Baltimore City undertook a study of 19 rail-highway grade crossings, with the objective of recommending improvements for those crossings. A portion of the study involved the development of a procedure that would rank the crossings in terms of their relative need for improvement. An equation that considered safety, vehicular delay, and emergency-access problem potential was used as the basis of the ranking procedure. The procedure also required the decision maker to quantify the relative importance of these three factors. The equation yielded a numerical score for each location under study. The crossings were then ranked on the basis of their scores. The crossing with the greatest score received the number 1 ranking; the rest of the crossings were ranked in descending order based on their scores. Although the procedure has several weaknesses (which are documented in this paper), it was found to be quite useful for the purposes of the study.

In 1982 the Interstate Division for Baltimore City (IDBC) undertook a study of 19 rail-highway grade crossings, with the objective of recommending improvements for those crossings. The crossings involved in the project had been identified by IDBC as being important, for one or more of the following reasons:

1. Under existing conditions, severe delays to roadway vehicles (especially emergency vehicles) resulted from extensive use of the crossing by trains,
2. An unusually high accident rate had been experienced at the crossing, and
3. Use of the crossing by trains had been projected to increase because of expanded coal export activities at the Port of Baltimore.

A portion of this project involved the development of a procedure that would rank the crossings in terms of their relative need for improvement. The purpose of this paper is to document the procedure thus developed.

RANKING PROCEDURE

The following general equation was used as the basis of the ranking procedure:

$$S_T = \sum_{i=1}^{n} W_i X_i \tag{1}$$

where

S_T = total score for a particular crossing,
X_i = normalized value of a parameter of concern, and
W_i = weighting given to parameter X_i, such that

$$\sum_{i=1}^{n} W_i = 1.00 \tag{2}$$

[Equations 1 and 2 were proposed by Davis (1) in 1979 as a technique for evaluating air pollution control strategies.]

The crossings were ranked on the basis of their total scores. The crossing with the greatest total score received the number 1 ranking, and the rest of the crossings were ranked in descending order based on their total scores.

The normalization of each of the parameters was accomplished by dividing all values of the parameter in question by the greatest value of that parameter and then multiplying the value thus obtained by 100. The resulting values were dimensionless and were between 0 and 100. This normalization was thought to be necessary to maintain a similar level in the contribution of each parameter to the total score. It was thought that use of a dimensionless term would be simpler than attempting to convert all parameters into the same units. The limitation that the weighting terms must add up to 1.00 forces the decision maker to rank the importance of a parameter relative to the other parameters.

For this project, it was believed that three separate elements should be considered in the ranking procedure:

1. The safety of motorists and pedestrians at the crossing,
2. The delay experienced by vehicles at the crossing, and
3. The potential for delay to emergency vehicles.

Thus Equation 1 became

$$S_T = (W_H)\,H + (W_D)\,D + (W_E)\,E \tag{3}$$

where

W_H = weighting given to the safety parameter of the motorist and pedestrian,
W_D = weighting given to the delay parameter,
W_E = weighting given to the parameter for emergency-access problem potential,
H = safety parameter of the motorist and pedestrian,

D = delay parameter, and
E = parameter for emergency-access problem potential.

The development of these terms is described in the following paragraph.

The question may be asked: Why bother rating crossings at all? Why not develop potential solutions for each crossing, and then perform cost-benefit analyses to rank improvements? Ideally, potential improvements for each crossing could be identified and developed to a level at which meaningful cost-benefit analyses could be performed. However, such a procedure could be quite lengthy and expensive, particularly if the number of crossings is large and if grade separations (the costs of which are quite site specific) are among the improvements to be considered. A priority rating system can quickly and relatively inexpensively identify those crossings that are most in need of improvement. Potential improvements for the short list of crossings can then be subjected to cost-benefit analyses at a fraction of the cost of performing cost-benefit analyses for potential improvements for all crossings.

DEVELOPMENT OF PARAMETERS

Safety Parameter

For simplicity, a hazard index was used as a measure of motorist and pedestrian safety at each crossing. (A hazard index is a numerical value obtained from one of several techniques.) The hazard index used for existing conditions in this project was the one currently used by the Maryland Department of Transportation (MDOT). Details regarding this hazard index may be found elsewhere (2). The hazard index used for future conditions was also one proposed for use by MDOT. Details regarding this hazard index may be found elsewhere (3).

The hazard index was normalized by dividing all values of the hazard index by the greatest value of the index and then multiplying all values thus obtained by 100. Thus all of the normalized values of the hazard index were between 0 and 100.

Delay Parameter

The delay parameter was developed as follows. Assume the existence of a hypothetical isolated rail-highway crossing at which a one-way, one-lane roadway is crossed by one railroad track. Let

λ = average roadway vehicle arrival rate at the crossing (vehicles/time),
μ = average roadway vehicle departure rate from a queue of stopped vehicles following a blockage of the crossing by a train (vehicles/time), and
T_S = time the crossing is blocked by a train.

Assume further that, at time T_O, a train starts to block the crossing. The length of the queue at the end of the blockage is given by

$$Q_{max} = \lambda T_S \quad (4)$$

where Q_{max} is the length of the queue at the end of the blockage (vehicles).

This situation is shown in Figure 1. The length of the queue at any time between T_O and T_R is the vertical distance from the time axis to the arrival line.

At time T_R the train completes its blockage of the crossing, and the queue of vehicles begins to discharge at a rate of μ. However, because vehicles continue to join the back of the queue at a rate of λ, the queue is not completely discharged in the amount of time given by Q_{max}/μ. Instead, as shown in Figure 1, the queue will be completely discharged at time T_F; that is, at the point where the arrival line and the departure line intersect. This point is shown as Point A.

The time required for the queue to clear completely (T_{QC}) may be found by determining the location of Point A. Assuming that the vertical axis passes through the time axis at T_R, Point A is defined by the following equations: $A = T_{QC}\mu$ and $A = Q_{max} + T_{QC}\lambda$. By substitution, $T_{QC}\mu = Q_{max} + T_{QC}\lambda$. Recalling that,

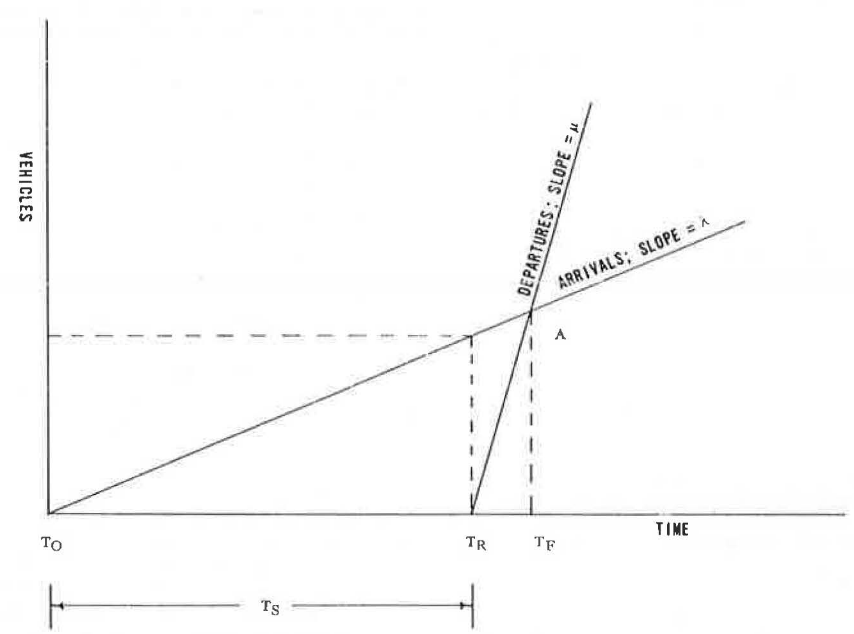


FIGURE 1 Roadway delay.

$$Q_{max} = \lambda T_S$$

$$T_{QC}\mu = \lambda T_S + T_{QC}\lambda$$

$$T_{QC}\mu - T_{QC}\lambda = \lambda T_S$$

$$T_{QC} = \lambda T_S/(\mu - \lambda) \quad (5)$$

The length of the queue at any time between T_R and T_F is given by the vertical distance between the arrival line and the departure line. At Point A, of course, this distance is zero.

The delay experienced by roadway vehicles as the result of a train crossing is equal to the area of triangle $T_O T_R A$. This area is given by

$$D = (1/2)\ (\text{base})\ (\text{height})$$

$$= (1/2)\ (T_S)\ [\lambda(T_S + T_{QC})]$$

$$= (1/2)\lambda T_S(T_S + T_{QC}) \quad (6)$$

where D is the total delay.

The delay experienced by roadway vehicles at any rail-highway grade crossing may now be found by relaxing the assumptions of the original model. For multiple-lane approaches, approach volume per lane is used to find D; D is then multiplied by the number of lanes. (This assumes, of course, that the approach volumes are equally divided among the approach lanes.)

If all trains are the same length, total vehicular delay may be found simply by multiplying the vehicular delay per crossing by the number of train crossings. If the trains are different lengths, however, the situation is somewhat more complicated.

A long train obviously blocks a crossing for a longer period of time than a short train does (assuming that the two trains travel at the same speed). Thus as the train length increases, T_S increases. As T_S increases, vehicular delay increases. However, as noted in Equation 7, delay does not increase linearly with T_S, but rather as the square of T_S. Recall Equation 6:

$$D = (1/2)\lambda T_S(T_S + T_{QC}) \quad (6)$$

Substituting for T_{QC} from Equation 5,

$$D = (1/2)\lambda T_S\ [T_S + (T_S\lambda/\mu - \lambda)]$$

$$= (1/2)\lambda T_S^2 + \{[(1/2)\lambda^2 T_S^2]/(\mu - \lambda)\}$$

$$= (1/2)\lambda T_S^2\ \{1 + [\lambda/(\mu - \lambda)]\} \quad (7)$$

The effect of this on the delay computations is that delay must be computed for each length of train and the results added together to obtain the total delay. It should be noted that this model has several weaknesses:

1. The model considers only peak-hour delay, and
2. The model is deterministic, and thus does not take into account variations in roadway and railroad traffic flow.

However, given the accuracy of the available information needed as input to the model, these limitations were not thought to be too severe.

The delay terms were normalized by dividing all values of the delay parameter by the greatest value of the parameter and then multiplying all values thus obtained by 100. Thus all of the normalized values of the delay parameter were between 0 and 100.

Parameter for Emergency Access Problem Potential

Emergency access problem potential was determined subjectively for each location on the basis of a consideration of the following factors:

1. Type of development in the general area of the crossing: The type of development was extremely important. Residential areas were believed to have a greater need for emergency access than did industrial areas (assuming that all other factors were equal) because of the concentration of population in residential areas.
2. Speed and convenience of alternate route: This factor also was extremely important because of the desirability of providing a convenient alternate route for an emergency vehicle that encounters a blocked crossing.

The emergency-access problem potential term was given a numerical value between 0 and 100, with 0 being lowest and 100 being greatest. There are two major disadvantages with this parameter: (a) the value of parameter is determined subjectively, and (b) the parameter requires thorough familiarity with the vicinity of each crossing and with the probable routes of emergency vehicles.

Weighting Factors

The weighting terms were determined entirely subjectively, subject to the constraint

$$W_H + W_D + W_E = 1.00,$$

as explained earlier. Obviously, the greater the weighting given to a term, the more important that term is. Once a weighting term was assigned a particular value, it retained that value for all the locations under study. (Of course, it would also be possible to allow the weighting factors to vary from location to location, as long as the summation constraint is met at each location.)

The use of the weighting factors requires an individual to quantitatively assess the relative importance of the various parameters. However, the subjective method of determining the weightings weakens the procedure. The weighting factors for the values are given in the following table:

Factor	Value
Delay (W_D)	0.25
Safety (W_H)	0.25
Emergency-access problem potential (W_E)	0.50
Total	1.00

USE AND RESULTS OF PROCEDURE

Two scenarios were evaluated in this project: existing conditions and future conditions following anticipated expansion of coal export activities at the Port of Baltimore. The raw values of the hazard index term, the delay term, and the emergency-access problem potential term under these scenarios are summarized in Table 1. The values in Table 1 were then normalized and placed on a worksheet, along with the weighting factors. The data in Tables 2 and 3 give the completed worksheets for each scenario.

Examination of the data in Tables 2 and 3 reveals that the rankings of the locations remain relatively constant. Waterview Avenue is ranked first under both scenarios, and rankings 2 through 5 include the same four locations, although the order changes. Only three locations had their rankings change by more than one position: going from existing conditions to future conditions, Bush Street went from ranking 10 to ranking 12; 2600 Hollins Ferry Road went from ranking 5 to ranking 2; and O'Donnell Street Service Drive went from ranking 14 to ranking 8.

Benhill Avenue received a high ranking (ranking 6 under existing conditions and ranking 7 under future conditions). This may at first be surprising, considering the low normalized delay value and the low normalized hazard index at this location. As the data in the tables indicate, however, the emergency-access problem potential value is quite high, and this is the major contributing term to this location's total score.

CONCLUSIONS

The ranking procedure developed in this project was found to be useful in ranking the crossings in terms of their relative need for improvement. The procedure offers the decision maker the opportunity to quantify the magnitude of the problems found at various locations, and also allows for the opportunity to quantify the relative importance of various problem parameters.

As noted in the paper, the procedure does have several weaknesses. The following specific shortcomings should be addressed in further study:

1. A technique for establishing the weightings should be developed.
2. An objective procedure for determining emergency-access problem potential should be developed. Such a procedure should consider quantitatively the length of the alternate route and the time required for use of that alternate route.
3. Monetarization of the parameters would be preferable to the normalization procedure used in this project.

ACKNOWLEDGMENT

The authors wish to thank K. Mammen Daniel for his assistance in the development of the vehicular delay model.

TABLE 1 Raw Values of Parameters

Location	Conditions	Hazard Index 1[a]	Hazard Index 2[b]	Peak-Hour Delay	Emergency-Access Problem Potential[c]
Warner Street	Existing	0.5476	0.2541	829.9	10
	Future	0.6013	0.2584	1,256.8	10
Ridgely Street	Existing	0.3452	0.0614	323.1	20
	Future	0.3776	0.0643	489.5	20
Bayard Street	Existing	0.2549	0.0477	89.2	50
	Future	0.2747	0.0502	135.1	50
Bush Street	Existing	0.3065	0.2252	278.1	50
	Future	0.3332	0.2288	421.0	50
2200 Hollins Ferry Road	Existing	0.1669	0.2103	99.3	100
	Future	0.2088	0.2180	197.5	100
2600 Hollins Ferry Road	Existing	0.3550	0.0587	194.7	85
	Future	0.5834	0.0735	584.7	85
2000 Hollins Ferry Road	Existing	0.1892	0.0457	135.8	100
	Future	0.2479	0.0519	432.5	100
Berlin Street	Existing	0.0870	0.1914	9.9	25
	Future	0.1058	0.1958	63.0	25
2000 Annapolis Road	Existing	0.2179	0.3838	182.4	40
	Future	0.2389	0.3876	317.3	40
2100 Annapolis Road	Existing	0.2210	0.3804	49.7	40
	Future	0.2561	0.3860	317.3	40
Kloman Street	Existing	0.1400	0.2072	158.2	60
	Future	0.1535	0.2099	275.6	60
Waterview Avenue	Existing	0.6048	0.0822	1,083.7	80
	Future	0.6898	0.0865	1,248.5	80
Benhill Avenue	Existing	0.1636	0.0387	9.6	90
	Future	0.2195	0.0482	74.8	90
Quarantine Road	Existing	0.0944	0.0268	3.3	85
	Future	0.2146	0.0493	101.4	85
Glidden Road	Existing	0.1199	0.0329	8.5	85
	Future	0.2519	0.0589	251.4	85
O'Donnell Street Service Drives	Existing	0.2383	0.0347	402.9	30
	Future	1.0524	0.0910	661.9	30
Newkirk Street	Existing	1.1207	3.8770	119.3	70
	Future	1.2051	3.9144	286.6	70
Ponca Street	Existing	0.2399	0.0509	118.9	70
	Future	0.2951	0.0573	381.1	70
Holabird Avenue	Existing	0.4185	1.2540	105.6	5
	Future	0.5165	1.2876	339.7	5

[a]Based on current MDOT technique; assumes all trains are day trains.

[b]Based on proposed MDOT technique; assumes all trains are day trains.

[c]Based on subjective weighting by project team.

TABLE 2 Location Ranking: Existing Conditions

Location	Weighting of Hazard Index (W_H)	Normalized Hazard Index (H)	Hazard Score: HS = $W_H \bullet H$	Weighting of Delay (W_D)	Normalized Delay (D)	Delay Score: DS = $W_D \bullet D$	Weighting of Emergency-Access Problem Potential (W_E)	Normalized Emergency-Access Problem Potential (E)	Emergency-Access Score: EAS = $W_E \bullet E$	Total Score: HS + DS + EAS	Project Ranking
Warner Street	0.25	48.9	12.2	0.25	76.6	19.2	0.50	10	5.0	36.4	11
Ridgely Street	0.25	30.8	7.7	0.25	29.8	7.5	0.50	20	10.0	25.2	17
Bayard Street	0.25	22.7	5.7	0.25	8.2	2.1	0.50	50	25.0	32.8	13
Bush Street	0.25	27.3	6.8	0.25	25.7	6.4	0.50	50	25.0	38.2	10
2200 Hollins Ferry Road	0.25	14.9	3.7	0.25	9.2	2.3	0.50	100	50.0	56.0	4
2600 Hollins Ferry Road	0.25	31.7	7.9	0.25	18.0	4.5	0.50	85	42.5	54.9	5
2000 Hollins Ferry Road	0.25	16.9	4.2	0.25	12.5	3.1	0.50	100	50.0	57.3	3
Berlin Street	0.25	7.8	2.0	0.25	0.9	0.2	0.50	25	12.5	14.7	18
2000 Annapolis Road	0.25	19.4	4.9	0.25	16.8	4.2	0.50	40	20.0	29.1	15
2100 Annapolis Road	0.25	19.7	4.9	0.25	4.6	1.2	0.50	40	20.0	26.1	16
Kloman Street	0.25	12.5	3.1	0.25	11.8	3.0	0.50	60	30.0	36.1	12
Waterview Avenue	0.25	54.0	13.5	0.25	100.0	25.0	0.50	80	40.0	78.5	1
Benhill Avenue	0.25	14.6	3.7	0.25	0.9	0.2	0.50	90	45.0	48.9	6
Quarantine Road	0.25	8.4	2.1	0.25	0.3	0.1	0.50	85	42.5	44.7	8
Glidden Road	0.25	10.7	2.7	0.25	0.8	0.2	0.50	85	42.5	45.4	7
O'Donnell Street Service Drive	0.25	21.3	5.3	0.25	37.2	9.3	0.50	30	15.0	29.6	14
Newkirk Street	0.25	100.0	25.0	0.25	11.0	2.8	0.50	70	35.0	62.8	2
Ponca Street	0.25	21.4	5.4	0.25	11.0	2.8	0.50	70	35.0	43.2	9
Holabird Avenue	0.25	37.3	9.3	0.25	9.7	2.4	0.50	5	2.5	14.2	19

TABLE 3 Location Ranking: Future Conditions

Location	Weighting of Hazard Index (W_H)	Normalized Hazard Index (H)	Hazard Score: HS = $W_H \bullet H$	Weighting of Delay (W_D)	Normalized Delay (D)	Delay Score: DS = $W_D \bullet D$	Weighting of Emergency-Access Problem Potential (W_E)	Normalized Emergency-Access Problem Potential (E)	Emergency-Access Score: EAS = $W_E \bullet E$	Total Score: HS + DS + EAS	Project Ranking
Warner Street	0.25	49.9	12.5	0.25	100.0	25.0	0.50	10	5.0	42.5	11
Ridgely Street	0.25	31.3	7.8	0.25	38.9	9.7	0.50	20	10.0	27.5	17
Bayard Street	0.25	22.8	5.7	0.25	10.7	2.7	0.50	50	25.0	33.4	14
Bush Street	0.25	27.6	6.9	0.25	33.5	8.4	0.50	50	25.0	40.3	12
2200 Hollins Ferry Road	0.25	17.3	4.3	0.25	15.7	3.9	0.50	100	50.0	58.2	5
2600 Hollins Ferry Road	0.25	48.4	12.1	0.25	46.5	11.6	0.50	85	42.5	66.2	2
2000 Hollins Ferry Road	0.25	20.6	5.2	0.25	34.4	8.6	0.50	100	50.0	63.8	4
Berlin Street	0.25	8.8	2.2	0.25	5.0	1.3	0.50	25	12.5	16.0	19
2000 Annapolis Road	0.25	19.8	5.0	0.25	25.2	6.3	0.50	40	20.0	31.3	16
2100 Annapolis Road	0.25	21.3	5.3	0.25	25.2	6.3	0.50	40	20.0	31.6	15
Kloman Street	0.25	12.7	3.2	0.25	21.9	5.5	0.50	60	30.0	38.7	13
Waterview Avenue	0.25	57.2	14.3	0.25	99.3	24.8	0.50	80	40.0	79.1	1
Benhill Avenue	0.25	18.2	4.6	0.25	6.0	1.5	0.50	90	45.0	51.1	7
Quarantine Road	0.25	17.8	4.5	0.25	8.1	2.0	0.50	85	42.5	49.0	9
Gildden Road	0.25	20.9	5.2	0.25	20.0	5.0	0.50	85	42.5	52.7	6
O'Donnell Street Service Drive	0.25	87.3	21.8	0.25	52.7	13.2	0.50	30	15.0	50.0	8
Newkirk Street	0.25	100.0	25.0	0.25	22.8	5.7	0.50	70	35.0	65.7	3
Ponca Street	0.25	24.5	6.1	0.25	30.3	7.6	0.50	70	35.0	48.7	10
Holabird Avenue	0.25	42.9	10.7	0.25	27.0	6.8	0.50	5	2.5	20.0	18

REFERENCES

1. C.F. Davis. A Framework for Assessment of Progress in Achieving Air-Quality Goals. Traffic Quarterly, Vol. 33, No. 2, April, 1979, pp. 263-274.
2. Rail-Highway Crossing Hazard Prediction Research Results. Report FRA-RRS-80-02. FRA, U.S. Department of Transportation, March 1980.
3. Rail-Highway Crossing Resource Allocation Procedure Users' Guide. Report FHWA-IP-82-7. FHWA, U.S. Department of Transportation, Dec. 1982.

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