TRANSPORTATION RESEARCH RECORD 681

Traffic Control Devices, Visibility, and Geometrics

TRANSPORTATION RESEARCH BOARD

COMMISSION ON SOCIOTECHNICAL SYSTEMS
NATIONAL RESEARCH COUNCIL

NATIONAL ACADEMY OF SCIENCES
WASHINGTON, D.C. 1978

Transportation Research Record 681
Price $5.40

mode
1 highway transportation

subject areas
21 facilities design
51 transportation safety
52 human factors
54 operations and traffic control

Transportation Research Board publications are available by ordering directly from the board. They may also be obtained on a regular basis through organizational or individual supporting membership in the board; members or library subscribers are eligible for substantial discounts. For further information, write to the Transportation Research Board, National Academy of Sciences, 2101 Constitution Avenue, N.W., Washington, DC 20418.

Notice

The papers in this Record have been reviewed by and accepted for publication by knowledgeable persons other than the authors according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

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

To eliminate a backlog of publications and to make possible earlier, more timely publication of reports given at its meetings, the Transportation Research Board has, for a trial period, adopted less stringent editorial standards for certain classes of published material. The new standards apply only to papers and reports that are clearly attributed to specific authors and that have been accepted for publication after committee review for technical content. Within broad limits, the syntax and style of the published version of these reports are those of the author(s).

The papers in this Record were treated according to the new standards.

Library of Congress Cataloging in Publication Data
National Research Council. Transportation Research Board.
Traffic control devices, visibility, and geometrics.

(Transportation research record; 681)

Seventeen reports prepared for the 57th annual meeting of the Transportation Research Board.

1. Traffic signs and signals–Congresses. 2. Visibility–Congresses. 3. Roads–Design–Congresses. 4. Traffic safety–Congresses.
I. Title. II. Series.
TE7.H5 no. 681 [TE228] 625.7'94 79-14725
ISBN 0-309-02826-4

Sponsorship of the Papers in This Transportation Research Record

GROUP 2–DESIGN AND CONSTRUCTION OF TRANSPORTATION FACILITIES
Eldon J. Yoder, Purdue University, chairman

General Design Section
Lester A. Herr, Federal Highway Administration, chairman

Committee on Geometric Design
B. H. Rottinghaus, Howard Needles Tammen and Bergendoff, chairman
Alvin R. Cowan, Federal Highway Administration, secretary
W. L. Brittle, Jr., Sumner B. Chansky, Robert R. Coleman, Harold D. Cooner, Hugh G. Downs, John C. Glennon, Malcolm D. Graham, C. William Gray, William R. Hawkins, Peter J. Hunt, Max N. Jensen, Milton L. Johnson, Richard D. Johnson, Frank J. Koepke, Jack E. Leisch, Alex E. Mansour, J. Robert Moore, Thomas E. Mulinazzi, Geoffrey M. Nairn, Jr., George B. Pilkington II, Stanley L. Ring, Robert A. Snowber, W. A. Wilson, Jr.

GROUP 3–OPERATION AND MAINTENANCE OF TRANSPORTATION FACILITIES
Adolf D. May, University of California, Berkeley, chairman

Facilities and Operations Section

Committee on Traffic Control Devices
Ken F. Kobetsky, West Virginia Department of Highways, chairman
H. Milton Heywood, Federal Highway Administration, secretary
John L. Barker, Robert L. Bleyl, Edward L. Cook, Charles E. Dare, J. R. Doughty, Roy D. Fonda, Paul H. Fowler, Robert M. Garrett, Alan T. Gonseth, Robert L. Gordon, Gerhart F. King, Kay K. Krekorian, Hugh W. McGee, Robert D. McMillen, Joseph A. Mickes, Zoltan A. Nemeth, Harold C. Rhudy, John C. Rice, David G. Snider, Ronald E. Stemmler, James I. Taylor, James A. Thompson, L. Thomas Van Vechten, Clinton A. Venable, Earl C. Williams, Jr., Walter P. Youngblood, Jason C. Yu

Committee on Visibility
Nathaniel H. Pulling, Liberty Mutual Research Center, chairman
Samuel P. Sturgis, Liberty Mutual Research Center, secretary
Albert Burg, Charles W. Craig, Warren H. Edman, Ralph A. Ehrhardt, Eugene Farber, W. S. Farrell, T. W. Forbes, Vincent P. Gallagher, S. A. Heenan, Robert L. Henderson, Robert T. Hennessy, Charles H. Kaehn, Antanas Ketvirtis, Lee Ellis King, Ken F. Kobetsky, Ralph R. Lau, Richard N. Schwab, Richard E. Stark, Frederick E. Vanosdall, Ned E. Walton, Earl C. Williams, Jr., William L. Williams, Henry L. Woltman, Kenneth Ziedman

Committee on Operational Effects of Geometrics
Stanley R. Byington, Federal Highway Administration, chairman
Robert B. Helland, Federal Highway Administration, secretary
Frank E. Barker, Robert E. Craven, Edwin W. Dayton, J. Glenn Ebersole, Jr., Daniel B. Fambro, Julie Anna Fee, John Feola, John C. Glennon, George F. Hagenauer, Rajendra Jain, Janis H. Lacis, William John Laubach, Jr., William A. McConnell, Woodrow L. Moore, Jr., Robert P. Morris, Thomas E. Mulinazzi, Sheldon Schumacher, James J. Schuster, Robert B. Shaw, C. Robert Shinham, Bob L. Smith, Robert C. Winans

K. B. Johns and Lawrence F. Spaine, Transportation Research Board staff

Sponsorship is indicated by a footnote at the end of each report. The organizational units and officers and members are as of December 31, 1977.

page 16, column 2, line 15 from bottom and lines 10-11 from bottom
Change "19.93 kN" to "2032 kg"
page 16, column 2, lines 5-6 from bottom
Change "19.88 kN" to "2028 kg"
page 18, column 1, lines 6 and 12
Change "10.05-kN" to "1025-kg" and "1.913-kN" to "195-kg"
page 18, column 1, line 20, and column 2, lines 20 and 23
Change "62.27-kN" to "6350-kg"
page 26, column 1, lines 13-15
Change each "nt·s" to "N·s" and each "lb-s" to "lbf·s"
page 26, column 2, lines 4-6
Change each "nt" to "N" and each "lb" to "lbf"
page 27, column 1, lines 1-2 and following table
Change to "production lot (1 kN = 225 lbf):

Test Piece	Axial Load (kN)
CT 7-11-1	136.3
CT 7-11-2	115.6
CT 7-11-3	116.1
CT 7-12-3	119.7
CT 7-12-4	122.1
Average	121.9"

page 27, column 1, lines 9-11 from bottom and following table
Change to "table (1 kN = 225 lbf):

Test Piece	Shear Load (kN/coupling)
CT 7-11-4	28.0
CT 7-11-5	17.3
CT 7-11-6	23.6
CT 7-12-1	17.3
CT 7-12-2	19.6
Average	21.2"

page 27, column 2, line 13
Change "227 kg (500 lb) and 4536 kg (10 000 lb)" to "2.2 kN (500 lbf) and 44.5 kN (10 000 lbf)"
page 28, column 1, lines 3, 14-15, 18, and 22
Change each "nt·s" to N·s" and each "lb·s" to "lbf·s"
page 29, Abstract, line 15
Change "362 kg·s" to "3.6 kN·s"
page 30, column 2, line 21
Change "1145, 1105, and 1060 kg·s" to "11.2, 10.8, and 10.4 kN·s"
page 30, column 2, line 17 from bottom
Change "492, 487, 500, and 464 kg·s" to "4.93, 4.89, 5.02, and 4.65 kN·s"
page 31, column 1, line 11
Change "350, 360, 350, and 357 kg·s" to "3.43, 3.53, 3.43, and 3.50 kN·s"
page 31, column 1, line 39
Change "338, 349, and 388 kg·s" to "3.31, 3.43, and 3.80 kN·s"
page 31, column 2, line 35
Change "91 kg·s" to "0.89 kN·s"
page 31, column 2, line 44
Change "Tunnel momentum change, kg·s 504 351 338 452" to "Tunnel momentum change, kN·s 4.93 3.43 3.31 4.43"
page 31, Table 3
Change the momentum change values from kg·s to kN·s for each category: "Speed Trap Measurement: NM, 4.75, 3.51,—,4.96"; "Integration of Tunnel Acceleration: 11.17, 4.93, 3.42, 3.31, 4.50"; "Integration of Rear-Deck Acceleration: 10.8, 4.89, 3.53, 3.43, 4.10"; "High-Speed Film Analysis: 10.4, 4.89, 3.43, 3.80, 4.48"
page 32, column 1, lines 7 and 10
Change "91 kg·s" to "0.89 kN·s" and "457, 418, 457, and 506 kg·s" to "4.43, 4.11, 4.25, and 4.97 kN·s"
page 33, column 1, lines 7-8
Change "500 kg·s" to "4.89 kN·s" and "350 kg·s" to "3.34 kN·s"
page 33, column 1, text table
Change the momentum change values from kg·s to kN·s for each test: "Test 1:—, 11.22, 10.83, 10.39"; "Test 2: 4.75, 4.94, 4.89, 5.02"; "Test 3: 3.51, 3.44, 3.53, 3.43"; Test 4:—, 3.31, 3.42, 3.80"; "Test 5: 4.96, 4.48, 4.10, 4.48"
page 33, column 1, line 31
Change "350 kg·s" to "3.34 kN·s"
page 33, column 2, lines 5 and 7
Change "500 kg·s" to "4.89 kN·s" and "91 kg·s" to "0.89 kN·s"

Transportation Research Record 681
page 19, column 2, line 22
Change "frequently" to "infrequently"

Transportation Research Record 720
page ii, Library of Congress data
Change "[666'.89]" to "[66';893]"

Transportation Research Record 721
page ii, Library of Congress data
Add "National Research Council. Transportation Research Board.
Rail and motor carrier reports.

(Transportation research record; 721)

1. Railroads—Addresses, essays, lectures. 2. Railroads—United States—Addresses, essays, lectures. 3. Transportation, Automotive—Freight—Addresses, essays, lectures. 4. Transportation—Law and legislation—United States—Addresses, essays, lectures. I. Title. II. Series.
TE7.H5 no. 721 [HE1031] 380.5s [385]
ISBN 0-309-02971-6 79-607922

Transportation Research Record 734
page ii, column 1
Change publication data to

Transportation Research Record 734
Price $3.40
Edited for TRB by Mary McLaughlin

mode
1 highway transportation

subject areas
24 pavement design and performance
31 bituminous materials and mixes
34 general materials
35 mineral aggregates
62 soil foundations

Change Library of Congress data to

Library of Congress Cataloging in Publication Data
National Research Council. Transportation Research Board.
Copper mill tailings, incinerator residue, low-quality aggregate characteristics and energy savings in construction.

(Transportation research record; 734)

Contents

Evaluation of Delineation Systems for the New Jersey Barrier

William L. Mullowney, New Jersey Department of Transportation

A prototype delineation system developed for installation on concrete median barriers is described. The visibility of the reflective devices used was evaluated with respect to the following factors: the effect of weathering on the reflectivity of the reflectors; the effect of weathering and other destructive forces on the durability of the reflectors; and the effects of vertical placement, opposing headlight glare, and wet nighttime conditions on the visibility of an installation. Mounting materials and techniques were evaluated to determine those that were most durable with respect to weathering and other destructive forces. To document permanently the effects of headlight glare and wet nighttime conditions on the visibility of the system, 8-mm color motion pictures were taken of an experimental installation.

The New Jersey type of concrete median barrier, known as a center barrier, has proved to be an effective countermeasure in head-on collisions. Although the barrier offers reduced accident severity to motorists, it may create a visibility problem at night for some drivers. In 1975, 258 single-automobile accidents that involved striking the center barrier occurred on 112 km (70 miles) of US-1 in New Jersey—135 at night and 52 under wet nighttime conditions. It is likely that other single-automobile strikings of the barrier go unreported, especially at night, since no other vehicle is involved and since the purpose of the median barrier is to redirect a colliding vehicle back into its own lane of travel.

Single-automobile center-barrier accidents result from what Alexander and Lunenfeld (1) describe as a "catastrophic system failure" of the "guidance level of driver performance." This performance level refers to the "drivers' task of selecting a safe speed and path on the highway." This selection involves evaluating the immediate situation, making appropriate speed and path decisions, and translating these decisions into vehicle-control actions. To perform these functions, the motorist needs to be provided with a sufficient number of unambiguous messages that are functional under a variety of weather conditions.

Delineation of median barriers will provide motorists with two guidance inputs to aid safe passage along the road. Immediately in front of the vehicle, such delineation will show where not to drive; that is, the median barrier will be perceived as a fixed, continuous, physical object to be aware of and avoided. Farther ahead of the vehicle, the reflectors will provide positive delineation by outlining the path of the barrier.

The necessity for delineation of median barriers is evident during nighttime driving conditions and especially during wet nighttime conditions. The visual contrast between the barrier and the roadway that supplies near and advance guidance information during daylight conditions is reduced during dry nighttime conditions and vanishes almost altogether in wet nighttime situations. The addition of a white pigment to the molded concrete has increased the contrast between the barrier and the road surface at night but is ineffective on wet nights. Delineators are needed to give the barrier a line of discrete visual cues that would replace or supplement the greatly diminished guidance information that exists under wet nighttime conditions.

STUDY DESIGN

The purpose of this study was to develop and test a delineator system that performs adequately on the median barrier after years of weathering. The characteristics that would affect the adequacy of the system were the visibility of the total system and the durability of its various parts.

Experimental variables were chosen for study if they were thought to affect the visibility or durability of the system. To study these variables, environmental factors that affected delineator performance were identified. The relation between these factors and the experimental variables was observed by means of performance measures developed and used during the study.

EXPERIMENTAL VARIABLES

Types of Reflective Devices

Various types of reflective devices were obtained from an extensive survey. The list of materials was narrowed down to six amber devices used in the major evaluations by using the relative reflectivity of new devices and adaptability to barrier application as acceptance criteria. The following devices were selected for the major tests (Figure 1):

Reflector	Type	Trade Name
1	Vinyl microscopic cube corner	Reflexite
2	Acrylic encapsulated lens sheeting	3M BD-21
3	Acrylic cube corner	Stimsonite 975
4	Silvered convex glass lens	Swareflex 3290
5	Wide-angle silvered acrylic cube corner	Stimsonite 2400
6	Low-profile acrylic cube corner	Stimsonite 960

Vertical Position on the Barrier

Three vertical positions were investigated during the project: on top of the barrier, on the side of the barrier 12.7 cm (5 in) from the top, and on the side of the barrier 35.6 cm (14 in) from the top. Originally, it was thought that headlight glare would render only the top-mounted devices ineffective, and therefore more emphasis was initially placed on the side-mounted locations.

Mounting Materials and Techniques

Mounting brackets consisted of steel, ethylene vinyl acetate (EVA), and scrap rubber. The EVA mount shown in Figure 2 is 10.2 by 10.2 by 5 cm (4 by 4 by 2 in), with 0.38-cm (0.15-in) thickness and holes 0.96 cm (0.375 in) in diameter. The EVA and scrap rubber were expected to be superior because of their flexibility and reduced potential danger on impact. Attachment materials studied included concrete studs (Figure 3) and butyl adhesives.

The mounting techniques used consisted of combinations of the various brackets and attachment materials. Only concrete studs were used on metal brackets, either studs or butyl adhesive were used

Figure 1. Median-barrier reflective devices.

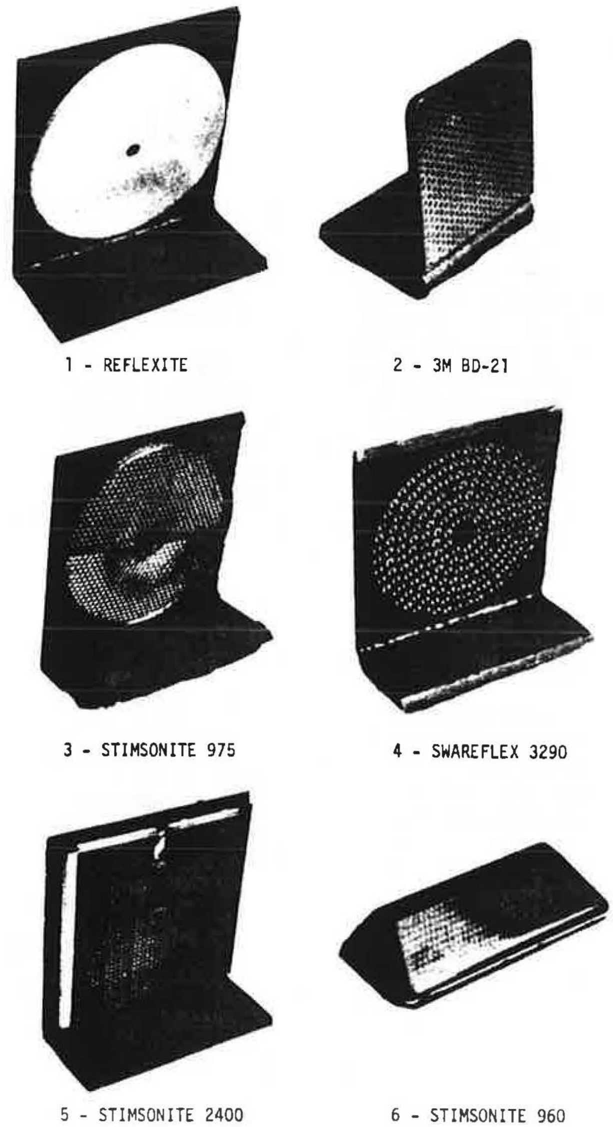


Figure 2. Ethylene vinyl acetate mount.

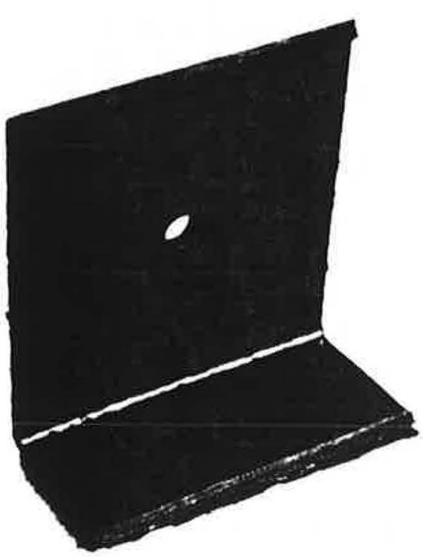


Figure 3. Concrete stud.

on EVA brackets, and only butyl adhesive was used on scrap-rubber brackets.

ENVIRONMENTAL FACTORS

Five environmental factors that could affect the visibility and durability of the experimental variables were chosen for study:

1. Weather conditions—The effect of rain on the reflectivity of the retroreflectors was considered important since the prime function of the system is to provide adequate visibility in inclement weather.
2. Dirt accumulation—It was expected that a layer of dirt on the surface of the reflectors would seriously degrade their reflectivity. How the individual devices were affected by this and whether any gross differences were discernible at the various vertical positions were considered to be important.
3. Wear from windblown particles—The scratching and pitting effect of windblown debris was monitored for the same reasons for which dirt accumulation was monitored.
4. Glare from opposing traffic—The effect of headlight glare on the visibility of the devices at the various vertical positions was studied.
5. Destructive forces—Whether any of the various reflector types or mounting materials or techniques were destroyed, lost, or rendered unusable was studied. Possible damaging forces were wet, plowed snow; impacts from vehicles or flying objects; and vandalism.

MEASURES OF PERFORMANCE

Dynamic Visibility Studies

The six test reflectors were mounted in groups on the barrier on US-1 in Trenton, New Jersey, so that the relative brightness of the individual devices at the vertical positions could be determined. A team of observers were to choose the brighter reflectors from a vehicle traveling in the left lane of traffic. The speed was about 64 km/h (40 mph), and low headlight beams were used. The team of observers consisted of engineers in the areas of traffic engineering, maintenance, quality control, and research. (The participating engineers' normal job responsibilities were related to delineation, but they were not familiar with this particular setup.) Groups of three or four raters were driven through the area and asked to fill out a questionnaire developed for the study.

Ratings were made (a) when the reflectors were new, (b) after one winter of weathering, and (c) after two winters of weathering (16 months of exposure on the barrier). The mounting configuration and the questionnaire were structured so that the following information was obtained:

1. The brightness rating of each device was compared against that of each of the others. The choices from all comparisons were totaled, and a final rating and a relative ranking were determined for each device.
2. The rater's direct preference for the various vertical positions was obtained by driving past a long stretch of the various reflectors at the different heights and considering them as a whole.
3. The rater's opinions on the adequacy of the devices as median-barrier delineators were obtained from consideration of each type of reflector over short stretches of highway.

Photometric Measurements

Specific intensity values of the six test reflectors were determined for three conditions: (a) when the reflectors were new, (b) after two winters of weathering when the reflectors were covered by dirt, and (c) after two winters of weathering when the reflectors had been cleaned. In addition, samples from each existing vertical position were removed and tested for the latter two conditions. All photometric tests were performed on an ESNA reflex photometer at incidence angles of 0° and 30° and divergence angles of 0.1°, 0.2°, and 0.5°.

Motion Pictures and Visual Observations

Eight-millimeter color motion pictures were taken of an installation over a long stretch of highway in both dry and wet nighttime conditions. A Kodak XL360 camera with Ektachrome ASA160, type G film was used. The driver and the camera operator also made visual observations of the effect of glare and the number of reflectors that could be seen in advance of the vehicle. These observations were later compared with similar observations taken from the developed film so that the reality of the motion pictures could be gauged. The film was used to allow all the staff engineers to review the installation under both wet and dry conditions.

Durability Survey for Mounts and Mounting Techniques

All analyses for durability were performed by visual observations. The various test locations were surveyed, and the devices were inspected for the following types of damage: permanent deformation of the bracket, looseness of the concrete bolts, rusting of bolts or rivets, missing reflectors or brackets, rusting of metal mounts, cracking of plastic mounts, and lifting and buckling of the butyl adhesive pads.

RESULTS

Effect on Reflectivity of Wet Nighttime Conditions

The reflectivity of devices used in an installation at New Brunswick, New Jersey, appeared to be enhanced during rain. This result was evident in both motion-picture analysis and visual observations made after 1 year of exposure. Project engineers reported that approximately five devices could be seen in advance of the automobile during dry nighttime conditions whereas 15 or more reflectors could be seen in the rain. Both observations were made while low headlight beams were being used. The visibility of the devices in the rain was limited by glare and geometry but not by reduced reflectivity.

The increase in the number of devices visible in the rain is thought to be caused by the following phenomena:

1. The rain may wash some of the dirt from the surface of the reflector and thus increase its reflectivity.
2. Decreased visibility of barriers and pavement markings may cause the barrier delineators to contrast more with the background.

Effect on Reflectivity of Dirt and Windblown Debris

As weathering or exposure time increased, the relative brightness of reflector 4 (convex glass lens) increased to the point that it was rated as the brightest after two winters of exposure (Table 1). (Ratings were calculated as follows: number of times selected as most reflective ÷ total number of comparisons with other reflectors.) This result was attributable to the dirt covering and the scratching and pitting from windblown particles observed on the surface of the reflectors. Documentation of this effect was also found in the photometric measurements. The glass reflector had a considerably smaller percentage reduction in specific intensity in all vertical positions when it was covered by dirt and when it was cleaned. The reduction in the photometric measurements after cleaning was caused by the scratching and pitting of the reflector surface by windblown particles.

The following percentages of original specific intensity for the six reflector devices resulted after two winters (16 months) of exposure at 0° incidence angle and 0.5° divergence angle:

Reflector	Top Side	Top Side Cleaned	Bottom Side	Bottom Side Cleaned
1	1	3	0	2
2	8	16	3	3
3	5	11	1	2
4	33	68	15	29
5	6	13	3	7
6	11	25	4	4

Some indication does exist that reflector 3 (the acrylic cube corner device) may retain superior reflectivity during rain. The results after one winter of weathering, given in Table 1, show that an acrylic cube corner received the highest rating when viewed in the rain and that reflector 4 (the convex glass lens device) was rated highest under dry conditions.

Adequacy of Retroreflectors as Median-Barrier Delineators

The raters viewed groups of reflectors at three vertical positions and determined whether they performed adequately as median-barrier delineators. A 50 percent threshold was chosen as a division between adequacy and inadequacy. The results indicated that after two winters of exposure all devices with the exception of the vinyl cube corner were considered adequate at the top and top-side positions. At the bottom-side position, one acrylic cube corner (reflector 3) and the convex glass lens (reflector 4) were judged adequate.

Effect of Dirt Accumulation and Windblown Particles at Various Vertical Positions

The first dynamic study, which rated unweathered reflectors, resulted in the bottom-side position being

Table 1. Comparison of retroreflectors in dynamic visibility studies.

Reflector	After Two Winters of Weathering (dry conditions)[a] Rating (%)	Rank	After One Winter of Weathering: Rain[b] Rating (%)	Rank	After One Winter of Weathering: Dry Conditions[c] Rating (%)	Rank	New Reflectors (dry conditions)[a] Rating (%)	Rank
1	16	6	0	6	0	6	42	2
2	28	5	6	5	23	5	17	6
3	57	2	95	1	75	2	78	1
4	77	1	37	3	86	1	40	3
5	33	3	28	4	32	4	34	4
6	32	4	41	2	43	3	22	5

[a]13 raters. [b]4 raters. [c]8 raters.

Table 2. Comparison of vertical positions of reflectors.

Position	Percentage of Comparisons in Which Position Was Chosen Most Reflective[a]: New	After One Winter	After Two Winters	Number of Times Position Was Chosen More Reflective[b]: New	After One Winter	After Two Winters
Top				0	10	13
Top side	30	48	47	2	0	0
Bottom side	44	32	24	8	0	0

Note: "Equal" judgments were not counted.
[a]Composite of results from seven locations comparing the six test devices.
[b]Results of single question concerning one location where a device was mounted at all three vertical positions.

selected as the most reflective. This was true not only for the reflectors judged collectively but also for each device individually. After exposure to the environment, the top-side and top positions were selected as the most reflective in all situations where they were used. In addition, as the exposure time increased, the trend toward higher ratings with increased height of mounting became more pronounced (Table 2).

These results can be attributed to the decreased amount of dirt covering and scratching and pitting experienced by the higher mounted devices. This effect is substantiated by the photometric data given previously, where the top-side position had a consistently smaller percentage reduction in specific intensity both when covered with dirt and when cleaned.

The following results, obtained from photometric evaluation of a location where the same device was mounted at all three vertical positions, proved informative:

Position	Reflectivity (percentage of original specific intensity after 30 months of exposure): Dirty	Cleaned
Top	31	47
Top side	3	6
Bottom side	2	3

The top-side position showed slightly less reduction in reflectivity than the bottom-side position; the top position was much less affected than the other two. It has been hypothesized that dirt and debris channeled down the side of the barrier by natural wind or the slipstream wind of vehicles account for the much greater wear and dirt covering of the side-mounted reflectors.

Effect of Opposing Headlight Glare on Reflectors at Various Vertical Positions

In the study performed at the Trenton site, the raters were asked what effect headlight glare had on their ability to view the reflectors. In the first study, 10 raters said the top- or top-side-mounted devices were affected more by glare than the bottom-side ones. One rater noted an equal effect, and 2 did not respond. In the second study, 10 raters said the top and top-side positions were affected more than the bottom-side position, and 2 reported an equal effect. At this site, the traffic volume was very low and the glare effects were intermittent.

At the New Brunswick site, three researchers viewed the reflectors at the peak evening hour in both dry and wet conditions. The effect of glare here was more dramatic. Platoons of cars traveling in the opposing direction "washed out" long stretches of the reflectors. Although all reflectors were mounted at the top-side position and no evaluation could be made of the effect of glare on the other positions, the extreme "blacking out" of barrier visibility appears to preclude the effectiveness of barrier-mounted reflectors under such conditions.

Effect of Destructive Forces on Mounting Materials and Techniques

The most durable mounting technique found in the study was a butyl adhesive pad attached to a low-profile marker. In 16 months of exposure at the northbound Trenton site and 12 months of exposure at the New Brunswick site, none of the reflectors were found to be missing (Table 3). At the New Brunswick site, however, part of the butyl pad was lifting off the barrier under nine of the reflector mounts.

After 16 months of exposure, 2 percent of the mounts that used a flexible bracket (EVA) attached to the barrier with a concrete stud were missing at the northbound Trenton site. Several of the mounts, however, did not remain taut against the barrier, and the mounting bracket rotated around the stud, causing a loss of view of the reflective device.

Flexible mounts (EVA or scrap rubber) attached with a butyl adhesive had a higher rate of loss than flexible mounts with concrete studs. At the northbound Trenton site, 7 percent were missing, and 7 showed a lifting of the butyl pad. At New Brunswick, 21 percent of the mounts were missing, and the butyl pad was lifting on 43. At the southbound Trenton site, 11 percent were lost in the 30 months of exposure, and lifting was not investigated.

The metal mount attached with a concrete stud experienced the highest loss rate. At the northbound Trenton site, 53 percent were missing, and seven of the remaining mounts were bent after 16 months of exposure.

The reason for the high loss of metal mounts was thought to be their inflexibility when they are impacted by some object or force, such as a vehicle, a flying object, or wet snow from plowing operations. When hit, the metal mounts apparently suffered deformation of the L-shaped bracket or failure of the concrete stud in the concrete or both, which diminished their continued effectiveness. A possibility also exists that metal mounts that have fallen off their barriers may pose a danger to motorists if kicked up into the air by vehicle tires.

The flexible mounts do not pose the same danger to motorists as metal mounts since they are plastic or

Table 3. Durability of mounts and reflectors.

Site	Type of Mount	Mounts Installed	Mounts Missing Number	Mounts Missing Percentage	Mounts Bent	Mounts With Butyl Pad Lifting Off	Damaged Reflectors
Northbound Trenton after 16 months exposure	Metal with concrete stud	45	24	53.3	7	-	6
	EVA with concrete stud	47	1	2.1	0	-	0
	EVA with butyl adhesive	2	1	50.0	0	0	0
	Butyl adhesive	20	0	0.0	0	0	0
	Scrap rubber mount with butyl adhesive	12	0	0.0	0	7	0
Southbound Trenton after 30 months exposure	Scrap rubber mount with butyl adhesive	148	17	11.5	-	-	-
New Brunswick after 12 months exposure	EVA with butyl adhesive	75	16	21.3	0	43	0
	Butyl adhesive	57	0	0.0	0	9	0

scrap rubber and apparently remain longer on the barrier. The flexible mounts that use concrete studs appear to be more durable than those that use a butyl adhesive pad, but rotation of the bracket around the stud could be a problem. This may result from loosening of the stud or nut when a flexible bracket bends under impact and puts a stress on the attachment mechanism.

Improper installation technique, the stress put on a butyl adhesive pad during impact, and vandalism are thought to be responsible for failures of the butyl adhesive pad method. During installation, the primer must be dry before the mount is attached to the barrier. If the primer is not dry or if insufficient force is applied to the base of the bracket during mounting, premature failure may result. Mounting the L-shaped bracket toward or away from oncoming traffic may make a difference in the amount of buckling or lifting caused by impacts. It is not known whether contraction-expansion effects during freeze-thaw cycles cause any lifting of the butyl pad.

Vandalism was apparent in one area of the southbound Trenton test site. The scrap-rubber mounts suffered a higher attrition rate in an illuminated interchange area than anywhere else. It has been reported to project personnel that there is pedestrian traffic at this section of US-1 even though a safer path is available. One reflector was found dangling from the barrier as if a vandal stopped before completing the act. Vandalism is suspected since mounts that use a butyl pad can be removed from the barrier by a slow, steady force whereas the large, instantaneous force of a vehicle impact apparently temporarily flattens the flexible bracket but does not rip the pad off the barrier. This occurrence was noted in New Brunswick where an EVA-butyl adhesive mount remained on the barrier even though a force from an impacting vehicle ripped the reflector off and forced the rivet and washer through the mount. The hole through which the reflector was riveted to the bracket was enlarged and elongated, but the butyl pad and mount were otherwise unaffected.

DISCUSSION AND SUGGESTED RESEARCH

The visibility of the retroreflective devices was enhanced during wet nighttime situations. Whether this was due to increased reflectivity or increased contrast with the road is not known. In either case, the motorist is supplied with the guidance information needed to perceive the barrier hazard. However, the effectiveness of the delineators is diminished when there is opposing headlight glare. The erection of glare screens may be a solution to this problem. Research into the use of delineators on barriers topped by glare screens would be necessary since the screen may, like the barrier, channel dirt into the face of the reflector. Moving the reflector to the top of the screen may cause a reduction in the visibility of the delineators since headlight intensity may drop off rapidly with increased height. Cook (2) found that 1.2-m (4-ft) high mountings resulted in shorter detection distances than did heights of 0.75 m (2.5 ft)—the approximate height of both car headlights and the center barrier. The cost of maintaining a center-barrier installation over a more extended period of time also needs investigation. Included in such a study could be possible cleaning methods, determination of the effective life of delineators and mounts, and whether a delineator similar to the glass convex lens reflector could be less costly if manufactured in the United States (thus saving on the original installation costs).

After two winters (16 months) of exposure on the barrier, all retroreflectors would be adequate at the top position (with the exception of the vinyl cube corner). Future studies might determine whether this trend would continue, that is, whether many types of retroreflectors would remain straight at the top position and for how long. A continuance of this result might allow considerations other than initial brightness to be primary in choosing a retroreflector. Such other factors could be cost, vulnerability, and resistance to vandalism.

A study of the varying rates at which harder surface materials of reflectors are affected by the elements may be useful. The dynamic visibility study performed in this project indicates that vinyl surfaces are most quickly affected and glass surfaces least quickly. Acrylic surfaces fall in between. Whether this trend would continue as exposure time increased is not known.

Further research is also needed in developing a more durable and inexpensive mounting technique. As a result of this work, it has been recommended that a concrete stud and a butyl adhesive pad be used for mounting. Although this combination of attachment methods was not studied, it is recommended over methods that use two concrete studs, one concrete stud, or the butyl adhesive pad alone for the following reasons:

1. The butyl pad would protect the barrier surface from spalling where the mount was attached. Two studs alone would not do this.
2. The butyl pad would protect the concrete stud from rusting.
3. The butyl pad would prevent rotation of the bracket around the stud.
4. The use of the concrete stud would prevent failure of the system as a result of the butyl pad lifting off the barrier.
5. The use of the concrete stud would offer more resistance to vandalism.

Documentation of these possible advantages is necessary. In addition, whether or not a steel or aluminum rectangular plate covering the entire face of the bracket base would be necessary to prevent lifting around the edges should be studied.

The longitudinal spacing of reflectors in this study was 24.4 m (80 ft) on tangents and 12.2 m (40 ft) on curves. Increased spacing would certainly lower installation and maintenance costs, but what effect this would have on the overall effectiveness of an installation is not known. Shorter spacing would result in increased costs but might help combat the effect of glare. Shorter spacing may also be necessary in areas of extremely high dirt accumulation such as intersections. Research into these areas may prove helpful. It could be hypothesized that extremely bright reflective devices could in themselves cause a glare problem if they were spaced too closely. However, none of the products evaluated in this study were found to cause such a problem.

Whether a highly visible, durable center-barrier installation has any beneficial effects on road safety could be studied to further justify general use of such devices. Before-and-after accident analysis and other traffic performance measurements, such as lateral placements and lane volumes under wet nighttime conditions, might be used in this endeavor. The installation of center-barrier delineators along with reflective pavement markers meant to perform in inclement weather might have a beneficial effect.

ACKNOWLEDGMENTS

I wish to acknowledge the assistance provided by the following individuals and groups in completing this project: Arthur Roberts for his guidance and technical support in all phases of the project, including his critique and editing of this report; R. L. Hollinger and E. F. Reilly for their project direction and administrative assistance; Jeffrey Gertler and M. V. Jagannath for their work in coordinating much of the field work; the Bureau of Maintenance for the safety setups supplied by the Lawrence and College Farm Yards; the members of the Delineation Committee for their generous donation of time, technical evaluation, and ideas; and the various manufacturers who supplied the reflective and mounting materials used in the project.

The contents of this report reflect my views, and I am responsible for the facts and the accuracy of the data presented. The contents do not necessarily reflect the official views or policies of the New Jersey Department of Transportation or the Federal Highway Administration. This report does not constitute a standard, specification, or regulation.

REFERENCES

1. G. Alexander and H. Lunenfeld. Positive Guidance in Traffic Control. Federal Highway Administration, U.S. Department of Transportation, April 1975.
2. K. Cook. Reflector Analysis. Century Research Group, Arlington, VA, Aug. 1969, pp. 163-164.

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

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

Abridgment

Evaluation of Yellow-on-Brown Road Signs for the Adirondack Park

Gary F. Gurney, Earl D. McNaught, and James E. Bryden, New York State Department of Transportation

In 1892, a state park was established in northern New York. This 23 413-km² (9000-mile²) area, known as the Adirondack Park, is guaranteed by the state constitution to remain "forever wild." The Adirondack Highway Council, which is composed of several representatives of state agencies and the public, was convened in 1974 to formulate and implement a state policy of enhancing the appearance of park highways in the Adirondacks. In 1976 and 1977, the work of the council focused on the aesthetic appearance of road signs. They recommended that certain types of highway signs be colored yellow on brown rather than a standard white on green, blue, or brown. This combination was recommended because, over a 40-year period, these have come to be recognized as Adirondack Park colors. Thousands of brown wooden signs with yellow letters have been used throughout the park by the New York Environmental Conservation Department to identify camping areas, hiking trails, ski slopes, and other places of interest. In addition, because of a 1924 state law governing commercial signing, many private organizations and businesses in the park area have chosen to adopt these colors in their advertising.

Before a color change could be implemented, it was necessary to obtain a variance from nationally mandated signing standards. To obtain such a variance, it had to be shown that the new combination would perform as well as standard colors.

A review of existing literature showed several studies that related directly to the proposed research. Unfortunately, although each was complete within its own objectives, not enough information had been collected to answer our question: Would yellow on brown perform as well as standard color combinations for the general driving population? We also wished to survey the opinions of the motoring public on the proposed colors.

The study was divided into two phases: (a) aesthetic appraisal (both by photographic documentation and an opinion survey) and (b) measurements of visibility and legibility. This research is described in greater detail elsewhere (1).

AESTHETIC APPRAISAL

Investigation

A section of NY-28 and NY-30 between Indian Lake and Blue Mountain Lake in Hamilton County was reconstructed in 1976. In keeping with its designation as an environmental highway project, yellow-on-brown signs were installed along this highway on an experimental basis. These include minor destination signs (D 15 series), which are normally white on green; parking and rest area signs (D 30 series), which are normally white on blue; and miscellaneous information signs (D 61 series), which are also normally white on green. D 15 and D 30 signs were constructed of brown engineering-grade reflective sheeting with yellow engineering-grade cutout letters and applied to metal substrates. The D 61 signs were wooden with routed letters and painted yellow on brown.

This 18-km (11-mile) highway section contains a total of 33 yellow-on-brown signs in the three series listed. They range from 0.61 to 3.05 m (2 to 10 ft) wide and from 45.7 to 142.2 cm (18 to 56 in) high. Length of message varies from two to nine words. In addition, several small symbol signs identify hiking trails, snowmobile trails, and bicycle routes.

Because of the availability of these yellow-on-brown signs, the site's location in the heart of the park, and heavy tourist traffic during the summer, this section was ideal for the aesthetic appraisal, which consisted of photographic documentation and a driver opinion survey.

Photographic Documentation

For the photographic documentation, four yellow-on-brown signs were selected on the basis of sight distance and background: one D 15, two D 30s, and one D 61. Overlays constructed of engineering-grade reflective sheeting and letters in the standard colors of white on green or blue duplicated each yellow-on-brown sign. The actual signs and standard-color overlays were photographed under five background conditions: winter (snow), spring (primarily brown background), summer (green background), fall (multicolored foliage), and night.

The color photographs and movies provide a comparative record of the yellow-on-brown signs and the standard white on green and blue. As might be expected, the photographic colors do not precisely match those experienced by the human eye, but representation was adequate in most cases to provide a sense of how a particular sign color fits into the highway environment. This documentation was particularly useful in examining camouflaging by roadside vegetation or background colors similar to the sign.

No distinct advantages are detectable for any of the colors from these photographs. Generally, the green and blue signs stand out better under spring and fall conditions, which present a primarily brown background. In summer, the brown signs appear to stand out better against the primarily green background. During winter (white background) and at night (black background), all three colors stand out well.

Opinion Survey

The opinion survey was conducted during the week of July 11 to 15, 1977; interview stations were set up at Indian Lake and Blue Mountain Lake. One interviewer was positioned at each station to solicit verbal opinions from drivers. Surveys were conducted both during daylight hours and after dark. Half the drivers were "alerted"—i.e., stopped at the station before entering the test section and requested to observe the highway signs ahead and be prepared to answer questions at the second interview station. The other drivers were "unalerted"—i.e., given no information until they were questioned at the second station as they left the test section.

To guard against biasing the answers by the methods used by the two interviewers, the wording of each question was rehearsed beforehand to ensure uniformity between the interviewers. The following principal questions were asked about the signs:

1. Did you have any trouble spotting the yellow-and-brown highway signs?
2. Did you have any trouble reading them?
3. Compared with normal road signs, how well do you think the yellow-and-brown signs complement the Adirondack environment?
4. Would you like to see more widespread use of yellow-and-brown signs in the Adirondacks?

The unalerted drivers were asked several preliminary questions to determine if they had noticed the yellow-on-brown signs. Because some noticed a number of different items along the roadway, it was occasionally necessary to direct their recollection to these signs in particular. This did not appear to bias results because most drivers tended to give specific answers one way or the other. However, all indefinite responses were counted as negative responses. Several additional questions were asked to obtain a driver profile. Drivers were classified by sex, age (as estimated by the interviewer), home address, and frequency of use of this highway section.

The geographic distribution of sample drivers and the overall results of the opinion survey are shown in Figures 1 and 2 respectively. All differences in tabulated data and potential recorder bias were examined statistically. The results shown in Figure 2 are very favorable toward yellow-on-brown signs. Three-fourths of the drivers thought these signs complemented the environment better than white on green or blue, and four-fifths favored more widespread use of this color scheme in the Adirondack Park. About five out of six expressed no difficulty in spotting or reading these signs. In addition,

1. No difference of opinion was found between tourists and local drivers,
2. Night drivers were much more observant and slightly more critical of the signs than daytime drivers, and
3. No difference of opinion was found among groups stratified by age or sex.

LEGIBILITY AND VISIBILITY

A 15-km (9-mile) section of NY-9H in Columbia County was selected as a test site because of its rural nature, low traffic, and absence of roadside lighting. Its abundance of long tangent sections allowed long sight distances, and its closeness to Albany simplified the logistics of conducting a large-scale test.

Within this section, 18 test signs were erected at random locations along various tangent sections. They included six from each of the three sign series installed on NY-28 and NY-30. Nine signs were yellow on brown, six were white on green, and three were white on blue. The materials, duplicating actual signs in the Adirondack Park, consisted of engineering-grade reflective sheeting and letters on metal substrates. Again, the yellow-on-brown D 61 series signs were routed letters on wooden panels painted (not reflectorized) yellow and brown. Series D letters 15.24 cm (6 in) high were used throughout.

Each sign contained a nonsense message composed of words easily read but conveying no meaningful message to the reader. This type of message was used to ensure that the sign was read entirely and that the test subjects did not rely on glance recognition.

The subjects, employees of the main office of the New York State Department of Transportation (NYSDOT) in Albany, were screened in an effort to ensure that the sample would be representative of the normal driving population. Engineers and technicians involved in any phase of highway engineering were excluded. The visual acuity of each volunteer was tested by the NYSDOT Health Services Unit, which also checked for color blindness.

Sign legibility and visibility were measured under three sets of conditions—spring, summer, and night. (Winter measurements could not be obtained before the end of suitable snow cover in March.) Fifty subjects were tested in each group, and profiles for each group were balanced to the extent possible.

The test vehicle, a 1974 Matador station wagon, was equipped with a distance-measuring system capable of recording reactions of two test subjects simultaneously. To ensure valid results, specific detailed instructions were given to the subject before testing, and a series of practice measurements were made before reaching the test site.

Test Subject Profile

Profiles of test subjects were compiled in terms of driving experience, age, sex, and education. To the greatest extent possible, profiles were matched for each test group to guard against biasing results by the selection of sample characteristics. Subject profiles were also compared with the general driving population to ensure that results were valid on a general basis. Overall, the profiles appeared balanced and representative of the statewide driver population. Two significant exceptions were noted and considered during data analysis:

1. Because test subjects were solicited from within NYSDOT, the sample contained a large proportion of persons of working age.
2. Judging from the three parameters of annual distance driven, type of driving, and years of experience, it became apparent that relatively few inexperienced drivers were included in the sample.

Legibility and Visibility Distance

Visibility and legibility distances for the test signs are given in Table 1. In most cases, the yellow-on-brown signs could not be read from as far away as could white on green or blue. Although most differences were statistically significant, the absolute differences were small—11 percent in the extreme case. The average daytime legibility distance for the standard-color signs is 103 m (337 ft) compared with 99 m (325 ft) for yellow on brown. Traffic signs are commonly designed on the basis of 6 m of legibility distance for each centimeter of letter height (50 ft/in); this results in a legibility distance of 91 m (300 ft) for the test signs, which had 15.24-cm (6-in) letters. Both the standard and the special colors exceeded that value.

Visibility distances of yellow-on-brown signs were also slightly less, averaging 462 m (1515 ft) for all daytime readings compared with 493 m (1617 ft) for the

Figure 1. Geographic distribution of 313 sample drivers.

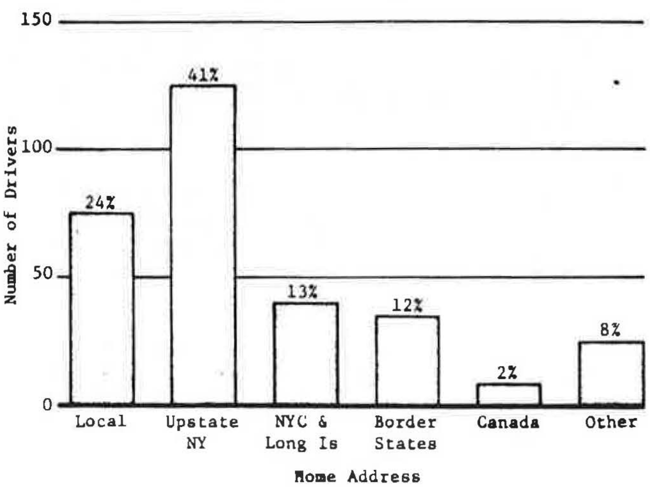


Figure 2. Summary of driver responses to opinion survey.

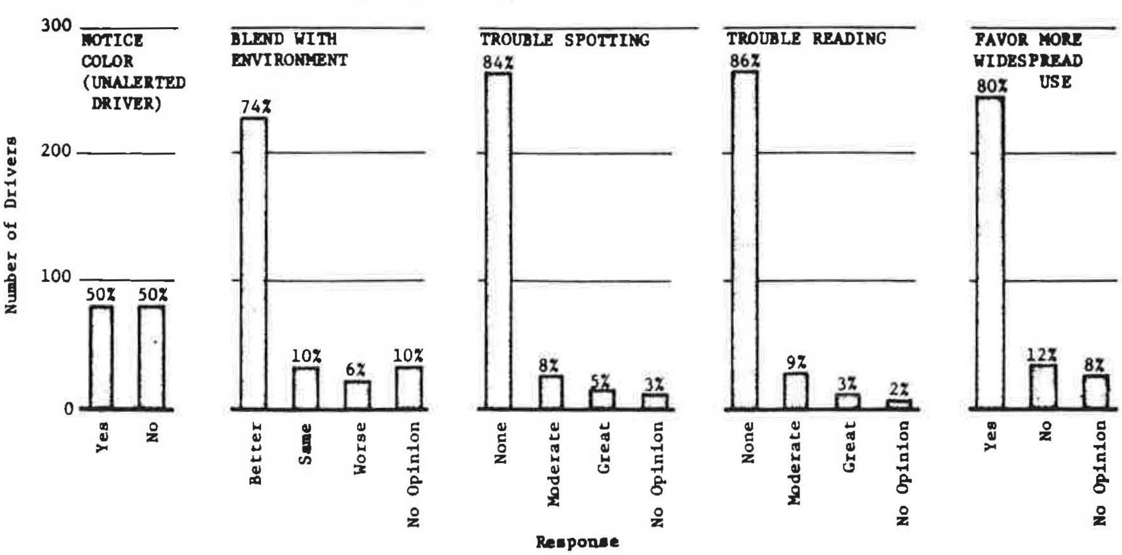


Table 1. Legibility and visibility distances for test signs.

		Spring			Summer			Night		
Item	Type of Sign	$\bar{X}$ (m)	s (m)	n	$\bar{X}$ (m)	s (m)	n	$\bar{X}$ (m)	s (m)	n
Legibility distance	Three-line blue	101	38	147	112	30	143	88	25	127
	Three-line brown	94	34	149	100	27	140	79	23	133
	Two-line green	98	36	148	106	27	147	80	23	133
	Two-line brown	92	33	147	103	28	142	77	24	131
	One-line green	93	32	148	107	27	142	81	23	128
	One-line brown	99	33	148	107	25	147	-[a]	-[a]	-[a]
Visibility distance	Three-line blue	593	186	135	511	136	141	413	132	134
	Three-line brown	472	167	142	422	170	142	399	99	135
	Two-line green	564	142	145	457	149	143	498	119	131
	Two-line brown	532	206	146	555	188	143	418	156	133
	One-line green	491	152	141	414	100	142	410	130	134
	One-line brown	477	150	130	445	133	130	-[a]	-[a]	-[a]

Note: 1 m = 3.3 ft.
[a]Not reflectorized.

standard colors. Again, although most differences in visibility distance in Table 1 are statistically significant, these small differences appear to have little practical meaning.

Specific Effects

Visibility and legibility distances were compiled with regard to the parameters of static visual acuity, color blindness, seating position, and background environment:

1. Generally, acuity correlated closely with both sign visibility and legibility. Unfortunately, we were not able to investigate the relation between aging and certain visual difficulties because of the relatively small sample of volunteers over age 60.
2. Differences appeared between color-blind subjects and the overall sample, but no particular problem is apparent for any one color.
3. Differences in readings could not be attributed to seating position for the daytime survey, but at night drivers were able to spot a sign more quickly than their passengers.
4. Spring readings were lower than summer readings, and night measurements were the lowest of the three surveys. Within each survey, however, the yellow-on-brown signs generally measured slightly lower than the standard-color signs.

CONCLUSIONS

1. All three sign-color combinations tested were legible beyond the accepted standard of 6 m/cm (50 ft/in) of letter height during daylight. This standard was published in 1939 (2) as a result of full-scale tests with black-on-white signs, and others (3, 4) have expanded on this initial research. Our study varied from its predecessors in that it combined (a) full-scale testing with (b) a relatively large number of subjects of various backgrounds by using (c) three different color combinations.
2. Differences in legibility and visibility among sign colors were small but statistically significant in some cases. Recent studies that have used colors other than black on white have reached this same conclusion (5, 6). But those studies were conducted primarily in the laboratory, and our study must be considered primarily a full-scale field test.
3. Driver parameters of age, sex, driving experience, and visual acuity could not be related to differences in performance among sign colors. It must be noted, however, that our sample included few inexperienced drivers and few over the age of 60.
4. Four-fifths of the drivers interviewed in an opinion survey in the Adirondack Park favored use of yellow-on-brown signs on park highways.
5. Color photographs confirm the importance of background color in sign visibility. Each color combination tested was more visible against some backgrounds than others.
6. More widespread use of yellow-on-brown information signs can enhance the parklike appearance of Adirondack Park highways with no loss in highway safety or motorist convenience.

ACKNOWLEDGMENT

This paper was prepared in cooperation with the Federal Highway Administration, U.S. Department of Transportation. In 1978, the Federal Highway Administration agreed to a variance permitting yellow-on-brown highway signs in the Adirondack Park.

REFERENCES

1. G. F. Gurney, E. D. McNaught, and J. E. Bryden. Evaluation of Adirondack Park Signs. Engineering Research and Development Bureau, New York State Department of Transportation, Albany, Research Rept. 56, Dec. 1977.
2. T. W. Forbes and R. S. Holmes. Legibility Distance of Highway Destination Signs in Relation to Letter Height, Letter Width, and Reflectorization. Proc., HRB, Vol. 19, 1939, pp. 321-335.
3. T. M. Allen. Night Legibility Distances of Signs. HRB, Bulletin 191, 1958, pp. 33-40.
4. D. E. Cleveland. Overhead Sign-Illumination Relationships. Texas Transportation Institute, Research Rept. 5-8, 1966.
5. T. W. Forbes and others. Luminance and Contrast Requirements for Legibility and Visibility of Highway Signs. TRB, Transportation Research Record 562, 1976, pp. 59-72.
6. B. L. Hills and K. D. Freeman. An Evaluation of Luminance Contrast Requirements in Fully Reflectorized Signs. Proc., Australian Road Research Board, Vol. 5, Pt. 3, 1970, pp. 67-94.

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

Abridgment

Comprehensive Evaluation of Nonsignalized Control at Low-Volume Intersections

Dana L. Hall,* Mitre Corporation, McLean, Virginia
Kumares C. Sinha and Harold L. Michael, Purdue University

A common practice by traffic authorities is to install stop signs at low-volume rural and urban intersections. This action generally is taken to ensure safety and in response to the lack of any clearly defined signing mandate. However, overuse of stop signs needlessly increases driver disobedience, travel time, and gasoline consumption. A recent research project conducted at Purdue University determined the most efficient signing policy for traffic flow through low-volume, unsignalized intersections.

DEVELOPMENT OF ANALYSIS TOOLS

The research examined the influence intersection conditions have on safety, travel time, fuel economy, and exhaust emissions. Low-volume flows necessarily precluded full reliance on field measurement of the many variable combinations. Computer algorithms were therefore used to aid in the study of emissions, fuel use, and travel time. Probability-of-conflict techniques, used in conjunction with accident records, supported the safety portion of the analysis.

Two properly validated simulation aids were required. One was a traffic model sufficiently detailed to reproduce accurately the flow characteristics of low-volume, unsignalized intersections. The second tool needed was a program that could process the traffic simulation output on an individual vehicle basis and estimate the gasoline consumption and resulting exhaust products. The traffic model selected was the Urban Traffic Control Simulation (UTCS-1S) model of the Federal Highway Administration (FHWA) (1). An appropriate aid to fulfill the second function was the Automotive Exhaust Emission Modal Analysis Model of the U.S. Environmental Protection Agency (EPA) (2).

The EPA model (which is calibrated in U.S. customary units of measurement) estimates grams per mile of four exhaust emission products: nitrogen oxides (NO_x), hydrocarbons (HC), carbon monoxide (CO), and carbon dioxide (CO_2). The EPA model, operating in close conformity with the microscopic level of the UTCS model, calculates these quantities by analyzing the unique velocity-acceleration pattern characteristic of an individual vehicle trajectory. Knowledge of the carbon-based products emitted also permits the gasoline quantity consumed to be calculated by use of a carbon balance relation. Appropriate modifications were made to the UTCS package to provide the EPA program with the model year of each vehicle under consideration as well as the corresponding time-velocity pattern.

Only those vehicles that traverse the minor (lower volume or controlled) street at an intersection are affected by the type of control implemented. The behavior of drivers on the major road was assumed not to be influenced by the sign type on the minor road. Automobiles on the minor road, on the other hand, were forced to slow or stop, which resulted in a substantial deviation from their preferred trajectory and a subsequent increase in fuel consumption, emissions, and travel time.

STUDY ANALYSES AND COMPARISONS

Fuel Consumption

Each hour of intersection traffic flow, simulated by the modified and validated UTCS-1S program, produced a set of time-velocity profiles equivalent in number to the total traffic volume on the minor road. An estimate of the gasoline consumption of each vehicle on that road was then calculated by using the carbon balance equation. Finally, averaging fuel use data within each hour produced one value representative of the combination of major-road volume, minor-road volume, and type of control peculiar to that cell. This derived mean approximated the amount of gasoline required by the average minor-road automobile to traverse a distance measured from 61 m (200 ft) upstream of an intersection to an exit point 61 m downstream, including any slowing, turning, or stopping.

Statistical tests indicated a highly significant difference between the average amounts of gasoline consumed by automobiles on the minor street as a function of the type of control implemented. A single automobile requires 0.026 liter (0.0068 gal) to traverse a stop control, 0.023 liter (0.0062 gal) to yield, and 0.021 liter (0.0055 gal) at an unsigned intersection. Considered on an individual vehicle basis, the difference in gasoline use between a restrictive control such as a stop sign and a less positive, rules-of-the-road approach appears inconsequential. Adopting a daily or annual perspective for that same single intersection changes this conclusion markedly, however. It can be shown, for example, that one minor street that carries a total volume of only 200 vehicles/d but is controlled by a stop sign requires 170.4 liters (45 gal) more gasoline per year than it would if controlled by a yield sign.

The energy implications inherent in various regional signing policies were extended to the state of Indiana. A procedure based on urban population and rural area was developed to derive an estimate of 120 000 unsignalized intersections across Indiana. The analyses indicated an annual potential savings of several million liters of gasoline given a signing policy that emphasizes yield signs and no sign control rather than stop signs at low-volume intersections that have adequate sight distance.

Exhaust Emissions

Velocity and model year data developed by UTCS-1S and input to the EPA model permitted statistical comparisons to be conducted on CO, HC, and NO_x pollutants. Primary attention was given the impact of the type of sign on the quantity of CO emitted by automobiles traversing the lower volume road. The important conclusion reached was that each successive step toward more positive, restrictive control causes a significant increase in the CO emitted by the average minor-road automobile. CO emissions created by an automobile traversing 122 m (400 ft) of observation area were 66 g/km (107 g/mile) at a stop, 59.6 g/km (96 g/mile) at a yield, and only 52 g/km (84 g/mile) given no sign control. Similar, although less pronounced, trends were found in the comparison of unburned HC; very little impact on NO_x was exhibited.

Travel-Time Delay

Travel times through intersections under various nonsignalized controls were computed from the velocity profiles output by the modified UTCS model. For the purposes of this analysis, delay was defined as the difference between the actual time required to traverse the 122-m (400-ft) observation area and the time that would have been needed if the automobile had maintained the velocity recorded when it first appeared in the observation area.

A highly significant difference in minor-street travel time or delay was proved for different types of controls. Approximately 4 s more travel time was required for the average vehicle that faced a stop rather than a yield sign and over 5 s more by a vehicle that faced a stop instead of no sign at all.

An idea of the average annual delay that can be expected at one intersection is shown in Figure 1 as a function of minor-road traffic volume and the type of control implemented. The graph shows, for example, that a minor road that carries 200 vehicles/d and has a stop sign will cause an average annual delay of 160 h. If the road has a yield sign, however, only about half that amount of time will be required, and if it is not signed at all—assuming sight distances warrant no sign—an average annual delay of about 60 h can be expected.

Figure 1. Effect of type of intersection control on annual delay to minor-road traffic.

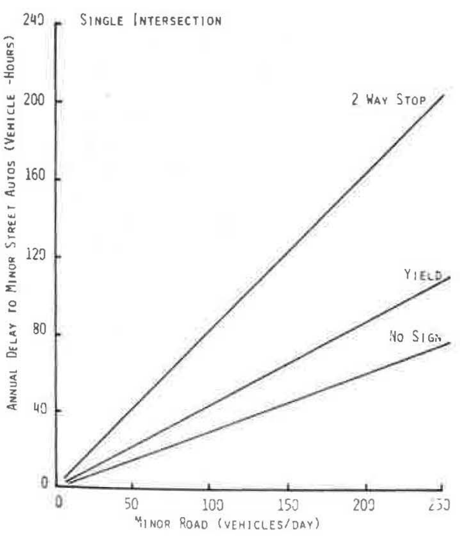


Table 1. Expected annual number of accidents by traffic volume and type of intersection control.

Control	Average Daily Traffic for Minor Road	Number of Accidents per Year — Average Daily Traffic for Major Road: 100	200	300	400
No sign	100	0.087	0.174	0.259	0.345
	200		0.345	0.516	0.686
	300			0.772	1.026
	400				1.363
Yield sign	100	0.052	0.104	0.155	0.207
	200		0.207	0.310	0.412
	300			0.463	0.616
	400				0.818
Stop sign	100	0.044	0.087	0.130	0.173
	200		0.173	0.258	0.343
	300			0.386	0.513
	400				0.682

Safety and Accidents

In support of the hypothesis that more efficient traffic flow can be attained by proper application of STOP and CROSS ROAD signs, Stockton, Mounce, and Walton (3) performed a comparison of two-way-stop and unsigned intersections based on probability-of-conflict concepts as well as accident and operating costs. That effort did not consider the effects of yield signs, but knowledge of accident reduction attributable to yield signs compared with no sign made it possible to incorporate all three control techniques.

Perkins (4) has shown the ratio of accidents to conflicts to be 0.000 33. Using that estimate and the conflict values computed by probability analysis yields the expected number of accidents per year at an unsigned road crossing. The literature search indicated that yield signs rather than no signs may reduce accidents anywhere from 20 to 60 percent. Therefore, an average accident reduction of 40 percent was assumed for yield signs in comparison with no control.

Contrary to common opinion, available literature based on accident records indicated that using a stop sign rather than a yield control had little effect on accidents. But the intent in this analysis to examine the stop sign in the best possible light permitted the assumption of a 10 percent accident reduction for stop signs in comparison with yield signs. Applying the 40 percent reduction descriptive of yield control or the approximate 50 percent reduction for stop signs allowed the appropriate accident figures to be derived from the no-control values. Table 1 gives a summary of the resulting annual accident count as a function of volume and intersection control. It is clear that a very small number of accidents can be expected at the typical low-volume intersection regardless of the type of control installed.

SELECTION OF PROPER INTERSECTION CONTROL

It was apparent at this stage of the research that, given adequate sight distance, yield signs are the most desirable form of control at low-volume intersections. Yield signs provide the optimal trade-off between the safety factor and the variables of travel time, gasoline consumption, and exhaust emissions. That conclusion was further substantiated by performing a cost-benefit analysis.

The cost components can be illustrated by the following equation: Total annual cost = gasoline + other automobiles + delay + accidents. Dollar values based on an Indiana study conducted by Hejal and Michael (5) were assigned to accidents by type of severity. These costs, appropriately updated to 1975 values, were combined with intersection accident experience to provide an average accident cost per intersection of \$2242. Applying this average unit cost to the expected accident counts given in Table 1 provided possible savings resulting from the increased safety attributable to more

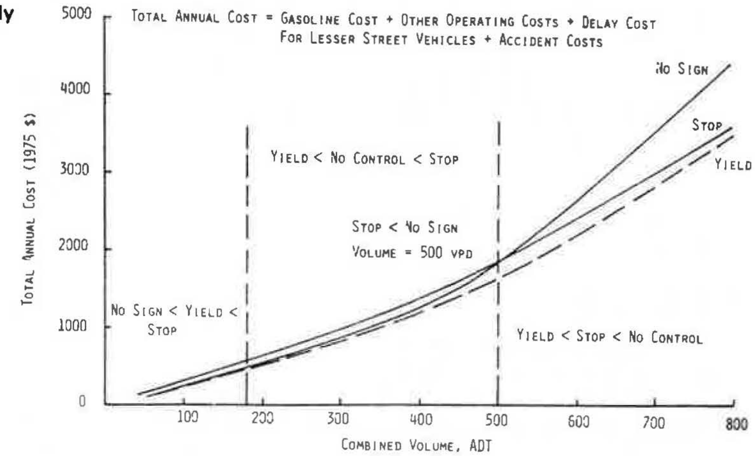


Figure 2. Expected annual cost per intersection for approximately equal split in traffic volume between crossing streets.

positive control at low-volume intersections.

Gasoline costs in 1975 were estimated at \$0.16/liter (\$0.60/gal) of which only \$0.13 (\$0.48) was actual cash outlay and \$0.03 (\$0.12) was refunded to the user through road-tax benefits. Other operating expenses include tires, oil, maintenance, and depreciation. These costs were estimated by updating the information given by Winfrey (6).

Delay costs were computed on the basis of a travel-time value study conducted by Thomas and Thompson (7). Using the census-estimated median income of Indiana families, given as \$9970/year, permitted the adoption of a set of the Thomas and Thompson travel-time values. The time values were prorated down to the average delay periods associated with stop, yield, and no control.

Figure 2 shows a graphical dollar trade-off between types of signs. It can be seen that, at total volumes from the upper limit [average daily traffic (ADT) of 800] of the low-volume crossings to roughly 200 ADT, the yield sign offers the lowest overall annual cost.

Yield signs provide a suitable compromise between the minimum operating cost of no signs and the minimum accident costs of stop control (this excludes consideration of the expense of sign installation and maintenance). Including installation and maintenance costs would show no sign at all to be the least expensive control at very low traffic volumes—perhaps intersection volumes in the range of less than 200 total vehicles/d.

ACKNOWLEDGMENT

This research was sponsored by the Highway Extension and Research Project for Indiana Counties at Purdue University. We are solely responsible for the results of the study.

REFERENCES

1. Application of Network Simulation Models to the Analysis of Urban Intersection Performance. Federal Highway Administration, U.S. Department of Transportation, Rept. FHWA-RD-74-25, Sept. 1973.
2. Automobile Exhaust Emission Modal Analysis Model. Office of Air and Water Programs, U.S. Environmental Protection Agency, Rept. EPA-460/3-74-024, Oct. 1974.
3. W. R. Stockton, J. M. Mounce, and N. E. Walton. Guidelines for Application of Selected Signs and Markings on Low-Volume Rural Roads. TRB, Transportation Research Record 597, 1976, pp. 26-32.
4. S. R. Perkins and J. I. Harris. Traffic Conflict Characteristics—Accident Potential at Intersections. HRB, Highway Research Record 225, 1968, pp. 35-43.
5. S. S. Hejal and H. L. Michael. The Estimated Direct Costs (1970) of Traffic Accidents on Indiana Rural State Highways. Purdue Univ., Rept. JHRP-20, 1970.
6. R. Winfrey. Economic Analysis for Highways. International Textbook Co., 1969, Tables A-8 and A-28.
7. T. C. Thomas and G. I. Thompson. Value of Time Saved by Trip Purpose. HRB, Highway Research Record 369, 1971, pp. 104-114.

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

**D. L. Hall was with Purdue University when this research was performed.*

Abridgment

Signing Warrants for Nonservice Facilities

William L. Mullowney, New Jersey Department of Transportation

The need for proper and sufficient right-of-way signing is apparent to all motorists. However, what constitutes proper and sufficient signing is not fully stated in the laws and regulations that cover this field.

Title 23 of the U.S. Code states that signing must give information "in the interest of the traveling public" and further that signing must "promote the highway's safe and efficient utilization." Title 27 of New Jersey state law requires the development of programs that "foster efficient and economical transportation services in the state," and Title 39 prohibits any commercial advertising on highways.

The Manual on Uniform Traffic Control Devices (MUTCD) (1) provides guidelines for regulatory, warning, and guide signs. The guide signs covered generally fall in two areas: (a) route information and (b) service facilities such as gasoline, lodging, telephone, camping, food, and hospitals. The entire area of nonservice facilities—that is, facilities that do not provide services necessary to the motorist's continued travel on the roadway—is not covered in the manual. One exception to this exclusion is the criteria for recreational areas found in the MUTCD.

The number of nonservice facilities in a given area is often larger than a highway can reasonably accommodate without overburdening the motorist with information. To provide a system for determining which facilities warrant signing, a specific set of written criteria is necessary.

The purpose of this survey was to review and critique the state of the art of nonservice facility signing. It is hoped that this information will be of assistance to traffic engineers in developing the written warrants needed.

The survey was geared for obtaining information on the warrants used by the 50 states for nonservice facility signing on expressways and freeways. The tables presented represent returns from 49 of the states. The survey was sent out to the states in November 1975. The responses were received between January and April 1976 and reflect the state of the art at that time.

SURVEY QUESTIONNAIRE

Information on 14 types of nonservice facilities was requested on the questionnaire. They are

1. Amusement parks;
2. Zoological or botanical parks;
3. Stadium sports facilities;
4. Colleges and universities;
5. Historical, recreational, or cultural facilities;
6. Airports;
7. Industrial areas;
8. Central business districts (CBDs);
9. Bus terminals or stations;
10. Bus park-and-ride lots and car-pool lots;
11. Rail terminals or stations;
12. Toll-collecting highways, bridges, and ferries;
13. Health care facilities (other than hospitals); and
14. Government facilities.

The states were asked whether they signed for each facility on expressways and freeways and whether they had any written warrants or criteria for this function. It was requested that copies of warrants and a drawing or picture of the actual signs be returned with the questionnaire. The states were also asked whether any distinction was made between using actual or categorical names in the sign legend.

Table 1. Criteria groups used in written warrants for signing of nonservice facilities.

Criteria Group	Number of States Using Criteria Group	Facilities for Which Group Is Used
Location of facility relative to highway	24	All except industrial areas and park-and-ride or car-pool lots
Level of activity (attendance, enrollment, number of mass transit movements)	22	All except industrial areas; CBDs; park-and-ride or car-pool lots; toll-collecting highways, ferries, and bridges; and nonhospital health-care facilities
Size or capacity (seating, parking, total investment)	9	Amusement parks; zoological or botanical parks; stadium sports facilities; colleges and universities; historical, recreational, or cultural facilities; and temporary events
Traffic generation characteristics	10	Amusement parks; zoological or botanical parks; stadium sports facilities; historical, recreational, or cultural facilities; park-and-ride or car-pool lots; rail terminals or stations; and temporary events
Availability	13	Amusement parks; historical, recreational, or cultural facilities; and temporary events
Type of ownership (public or private)	4	Zoological and botanical parks; stadium sports facilities; historical, recreational, or cultural facilities; and airports
Official recognition	7	Amusement parks; colleges and universities; historical, recreational, or cultural facilities; and airports
Public services provided	8	Colleges and universities; historical, recreational, or cultural facilities; airports; and temporary events
Attraction value	2	Amusement parks and historical, recreational, or cultural facilities
Population of surrounding area	8	Zoological or botanical parks; stadium sports facilities; colleges and universities; historical, recreational, or cultural facilities; airports; CBDs; nonhospital health-care facilities; government facilities; and temporary events

Table 2. AASHO guideline criteria for signing of traffic generators.

Type of Generator	Specific Criterion	Major Metropolitan Areas (>250 000 population)	Urban Areas (50 000-250 000 population)	Rural Areas (<50 000 population)
Airports	Number of regularly scheduled (one-way) movements per day	40	30	20
	Kilometers	3.2	6.4	8
Colleges and universities	Number of students in full-time enrollment[a] (prime criterion)	10 000	8000	6000
	Number of off-street parking stalls	500	200	200
	Kilometers	4.8	6.4	8
Military bases	Number of employees or permanently assigned personnel (prime criterion)	5000	5000	5000
	Kilometers	4.8	6.4	8
Arenas, auditoriums, beaches, convention halls, dams, fairgrounds, lakes, national historical sites, national parks, recreation areas, stadiums, state parks	Seating capacity	5000	5000	5000
	Number of parking stalls	500	300	200
	Annual attendance[b] (prime criterion)			
State police stations	Kilometers	1.6	1.6	3.2
Toll highways and bridges	Direct access from exit and part of state highway system			
Business districts	Direct access and not more than 8 km from interchange			

Note: 1 km = 0.62 mile.

[a]As many as 4000 part time students on a 2 for 1 basis may be used in meeting this criterion; i.e., the maximum credit for part-time students shall be 2000.

[b]Two-hundred-thousand people plus 20 000/1.6 km of distance from freeways up to 8 km plus 300 000/1.6 km for each additional 1.6 km over 8 km for all population groups, only those days with 1000 or more attendance will be considered.

SUMMARY OF EXISTING WARRANTS

Table 1 gives the ten criteria groups into which the specific criteria were categorized and the number of states that supplied warrants in these groups. Although it was not specifically requested, many states included warrants for temporary signing. This information is included in the availability category in Table 1.

Table 2 was developed in July 1971 by the Operating Subcommittee on Traffic Engineering of the American Association of State Highway Officials (AASHO) as a guideline for traffic-generator signing. Four of the types of facilities included in the questionnaire are not considered in this table. They are industrial areas, rail terminals, bus terminals, and park-and-ride or car-pool lots.

Forty-six of the 49 states that sent replies to the questionnaire responded to the question concerning the distinction between the use of actual or categorical names. Twenty of those that answered said they did make such a distinction.

DISCUSSION OF RESULTS

Quantitative warrants should be established by each state for its own purposes and according to its own needs. Many of the states have apparently modified the AASHO chart to their own needs, altering the criteria as their own conditions dictate. Sixty-two percent of the warrants are in a modified form of the criteria presented in the AASHO chart. These modifications did not use all the criteria presented in the AASHO chart for a given facility but always included what the chart refers to as the prime criterion.

The large majority—85 percent—of the warrants summarized are of a specific or quantitative nature. The remaining 15 percent of the warrants are divided into two groups. Ten percent use vague terminology such as close proximity, reasonable distance, significant, frequent, and major. Such use of vague phrases in written warrants does not lend itself to decision making. Facilities denied signing under vague warrants might consider the decision arbitrary and contestable. Use of more specific and measurable criteria can help avoid this situation. The other 5 percent are in the official recognition group. This group is not composed of written warrants but lists the results of the application of unstated warrants leading to such recognition. If official recognition is to be used to allow signing for a facility, then the specific criterion that leads to the recognition should be included in the warrants.

Critique of Criteria Groups

Location Relative to the Highway

The distance of the facility from the highway where the sign is to be placed was the most widely used criterion found in the survey. Distances up to 40 km (25 miles) existed in some warrants, but generally a range from 1.6 to 16 km (1 to 10 miles) was used. Many warrants increased minimum attendance requirements as the distance from the highway increased. This group is important since without it the number of facilities eligible for signing might exceed the ability of the road and the motorist to make use of them.

Level of Activity

All the criteria in this group were specific and to the point whether they used numbers of people or numbers of mass transit movements. Attendance figures usually ran between 100 000 and 1 000 000/year, but some lower figures were used. Required enrollment in schools ranged between 600 and 10 000 full- and part-time students. Number of employees ranged from 2000 to 5000. The number of mass transit movements ranged from 10 to 60/d. All activity criteria varied with distance from the highway and population range of the surrounding area.

Size or Capacity

All seating capacity warrants found in this survey were for 5000 seats. Warrants for parking stalls ranged between 200 and 500 stalls. One state submitted a warrant that used total investment as the criterion for signing. A more pertinent criterion might be yearly operating costs or gate receipts. However, since activity and location criteria can adequately gauge the need for signing, monetary considerations seem unnecessary.

Traffic Generation

With the exception of vehicles per day or event, the criteria in this group are vague. The vehicle warrants ranged from 1000 to 3000 vehicles/d or event but were not used in many of the warrants. One problem with this type of criterion is that counting vehicles may not be as easy as counting attendance.

Availability

The availability category is generally used with temporary signing warrants where it has considerable applicability. The minimum amount of time a facility must be open ranges from 1 d to 6 months, and during this time the facility must satisfy other criteria. One state imposes the annual minimum warrants. More common is the need to satisfy decreased criteria in the areas of attendance and vehicles for the time the facility is open.

Ownership

Since the primary purpose of signing is to aid the motoring public and to promote safe travel, the use of ownership criteria in warrants seems irrelevant. If the avoidance of advertising for a privately owned facility is interpreted to include use of the name, then generic or categorical names can be used on the sign. Twenty states indicated that they make such a distinction or restrictive interpretation.

Public Services and Attraction Value

Even when use of criteria in these two groups is specific, it must be secondary to use of criteria for activity, location, and size. Public services and attraction value do not measure the need to sign a given facility because of generation of traffic or activity nor do these criteria have any bearing on the number of facilities in a given area. Moreover, the exclusion of a large generator because of its failure to provide such items as restrooms or an attractive environment may cause inefficient and unsafe use of highways because insufficient information about the generator is supplied to the traveling public.

Population Range of Surrounding Area

Population range was the second most widely used criterion found in the survey; the states defined rural, urban, and major metropolitan areas according to their needs. As the population of the area increases, a facility generally must be closer and larger and must attract more people and vehicles to be eligible for signing. It is apparently thought that more populated areas have a greater density of nonservice facilities, which necessitates more stringent warrants to avoid placing excessive numbers of signs.

Placement and Number of Signs

When a large number of different types of nonservice facilities in a given area meet the criteria for signing, further warrants should exist to limit the number of signs placed. Three warrants for such limitations were presented by some of the states:

1. If there are more than a given number of qualifying generators, those that create the greatest traffic demand will be signed for.
2. The facilities that exceed the annual criteria by the largest percentage will be eligible for signing.
3. At a single exit where several generators qualify for signing, the following priorities shall prevail: road names, military reservations, cities, national parks, state parks, 4-year colleges, 2-year colleges, vocational and technical schools, and tourist attractions.

These limitations included warranting the number of signs allowed to be placed at a given intersection or exit. Generally, signs for two generators are allowed; one state indicated that signs for as many as four could be placed.

The first warrant stated above is not specific. Whether greatest traffic demand means vehicles per day or per event is unclear. The second warrant is specific but could lead to conflicting interpretations if two or more criteria are used to indicate eligibility. If this were the case, one facility might exceed another according to one criterion but not according to others. Distinguishing one of the criteria as the prime criterion, as suggested by AASHO, might eliminate potential problems. The third warrant should be expanded to include more types of generators such as transit facilities, shopping centers, business districts, airports, and government buildings. Tourist attraction could cover too many types of facilities and should be more specifically defined.

RECOMMENDATIONS

As far as possible, warrants should be composed of specific criteria. Most of the warrants summarized in this survey are of a quantitative or specific construction, yet vague, undefinable phrases are also used. Vague warrants offer little assistance in distinguishing borderline cases and leave the application of the warrant open to criticism.

In developing warrants, the major items of concern should be location relative to the highway, level of activity, size and capacity, and population of the surrounding area. The AASHO chart concisely combines these four groups into clear, quantitative warrants. Warrants of a similar form but tailored to the needs of the states might be considered for use. The other groups, with the possible exception of availability, appear to be extraneous and do not pinpoint generators whose signing would promote efficient highway travel.

It may also be important for any written warrant to contain an escape clause—i.e., a statement that can be used to deny signing for an otherwise eligible facility. However, the reasons for invoking this clause would have to be clear and reasonable and show that placement of a particular sign would have a negative effect on the promotion of safe and efficient highway travel.

ACKNOWLEDGMENTS

I wish gratefully to acknowledge the assistance and guidance provided by Arthur Roberts in the writing and editing of this paper. The cooperation of the states that supplied the information used in this paper is also acknowledged.

REFERENCE

1. Manual on Uniform Traffic Control Devices for Streets and Highways. Federal Highway Administration, U.S. Department of Transportation, 1971.

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

Abridgment

Performance of Signs Under Dew and Frost Conditions

John W. Hutchinson, University of Kentucky
Ted A. Pullen, Kentucky Department of Transportation

The nighttime legibility and target value of retroreflective traffic-sign legend and background materials are frequently decreased and occasionally lost because of dew or frost formation on the face of the sign. Dew forms when the temperature of the sign face approaches the dew point of the surrounding air; frost forms when this temperature is below the freezing point of water. Certain atmospheric conditions are known to be favorable for dew formation: (a) a clear sky, (b) a still atmosphere, and (c) a supply of moisture in the air around the sign, i.e., high humidity. The frequency and duration of occurrence of dew formation therefore vary with such factors as climate, locale, season of the year, and atmospheric conditions.

Several different types of materials are commonly used as retroreflective surfaces on traffic signs to convey information to motorists under low-visibility conditions. Field experience has shown that formation of dew and frost on such materials occasionally reduces the effectiveness of signs to varying degrees for different types of materials and sometimes even totally obliterates the message under headlight illumination. The distinctive shapes of some of the more important regulatory and warnings signs (STOP, YIELD, CAUTION, RAILROAD CROSSING, SCHOOL) (1) help to overcome such temporary losses of sign-legend effectiveness and have even led to the suggestion of clearly distinctive shapes for other signs, such as a pennant shape for DO NOT PASS and an arrow shape for ONE WAY. However, even these distinctively shaped signs lose varying amounts of target value for the different types of materials when headlight illumination is scattered and diffused by droplets of dew or crystals of frost on the face of the sign. In the case of freeway guide signs, the decrease in the visibility of the legend sometimes represents total loss of the sign message.

Efforts to overcome the effects of this phenomenon have perhaps justifiably been given lower priority than many other more urgent problems that need research. In the experience of many traffic engineers, the fogging over (dew) or frosting over of sign messages occurs only rarely, only after the evening rush hour, and with fairly predictable regularity only during certain seasons of the year. However, recent research suggests that a lessening of the conspicuity (target value) and specificity (clearness of message) of traffic signs has adverse effects on driver behavior (2, 3, 4). Furthermore, commonly accepted engineering and psychological principles for transmitting information to drivers clearly demand as much uniform signing redundancy (distinctive shape, color, and message) and target value as can be maintained under any given weather conditions (5, 6).

In the absence of suitably energy-conservative means of otherwise overcoming the adverse effects of dew and frost on existing signs, the signing materials industry should be encouraged to develop materials that are less subject to these effects. In the meantime, it is expedient to evaluate existing materials and consider the use of the least affected combinations of signing materials currently available.

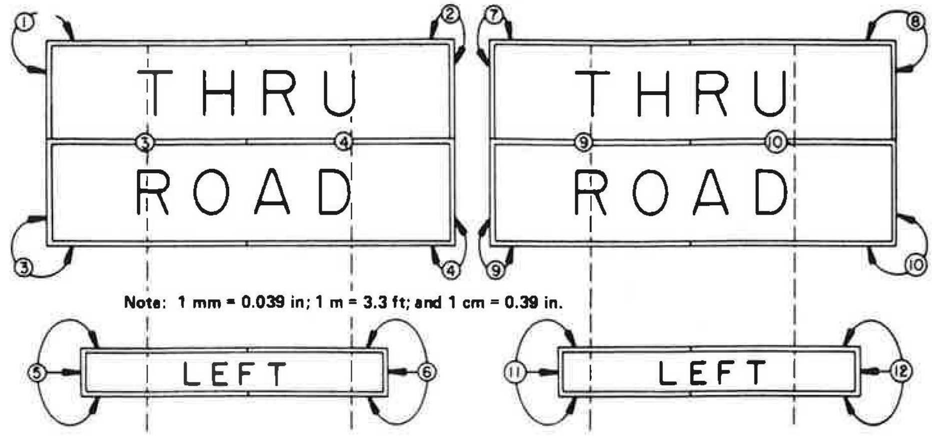
The legibility of signs under dew and frost conditions has been observed to vary with different combinations of legend, background, and mounting materials. The relative performances of different combinations of these materials have also been noted to vary somewhat with age (exposure) of the materials. The purpose of this study was to evaluate the effects of dew and frost on the nighttime legibility of several possible combinations of retroreflective legend and background materials under headlight illumination. Observations were made over an 8-month period, from April through November, in central Kentucky. A total of 31 nights with observed natural dew or frost formation between 9:00 p.m. and midnight were selected for purposes of sign evaluation.

TEST CONDITIONS AND MATERIALS

The observation site was at Blue Grass Field, the Lexington-Fayette County, Kentucky, airport, in a small valley surrounded on three sides by runways, taxiways, and airport service facilities. A National Oceanic and Atmospheric Administration weather service station at the airport provided ready access to needed atmospheric data.

Consideration of reported effects of light source and viewing conditions (7, 8, 9, 10) led to use of a standard automobile headlight system that was mounted on skids for mobility and maintained at constant brightness by a portable gasoline-powered generator and battery charger. Various combinations of encapsulated-lens, enclosed-lens, and button-copy materials were included in the test signs (Figure 1) (11, 12). All direct-applied legend and border materials were mounted in the shop. All demountable, embossed, and button-copy materials

Figure 1. Sign materials.



SIGN	LARGE 3.175 mm ALUMINUM SMALL 3.175 mm ALUMINUM	SIGN	LARGE 19.05 mm PLYWOOD, 1.58 mm ALUMINUM OVERLAY SMALL 3.175 mm ALUMINUM
BACKGROUND	TOP HALF ENCLOSED LENS BOTTOM HALF ENCAPSULATED LENS SMALL PANEL ENCLOSED LENS	BACKGROUND	TOP HALF ENCLOSED LENS BOTTOM HALF ENCAPSULATED LENS SMALL PANEL ENCAPSULATED LENS
COPY	TH ENCAPSULATED LENS DIRECT APPLIED RU ENCLOSED LENS DIRECT APPLIED RO BUTTON COPY AD ENCAPSULATED LENS DEMOUNTABLE LE ENCLOSED LENS EMBOSSED FT ENCAPSULATED LENS EMBOSSED	COPY	TH ENCAPSULATED LENS DIRECT APPLIED RU ENCLOSED LENS DIRECT APPLIED RO BUTTON COPY AD ENCAPSULATED LENS DEMOUNTABLE LE ENCAPSULATED LENS EMBOSSED FT ENCAPSULATED LENS DIRECT APPLIED
BORDER	① ENCAPSULATED LENS DIRECT APPLIED ② ENCLOSED LENS DIRECT APPLIED ③ BUTTON COPY ④ ENCAPSULATED LENS DEMOUNTABLE ⑤ ENCLOSED LENS EMBOSSED ⑥ ENCAPSULATED LENS EMBOSSED	BORDER	⑦ ENCAPSULATED LENS DIRECT APPLIED ⑧ ENCLOSED LENS DIRECT APPLIED ⑨ BUTTON COPY ⑩ ENCAPSULATED LENS DEMOUNTABLE ⑪ ENCAPSULATED LENS DEMOUNTABLE ⑫ ENCAPSULATED LENS DIRECT APPLIED

LARGE PANELS 1.22 x 2.44 m, COPY 38.1 cm (BUTTON COPY 40.6 cm), BORDER 2.54 cm
SMALL PANELS 1.83 x 0.305 m, COPY 15.2 cm, BORDER 2.54 cm

SIGN HEIGHT TO TOP 3.96 m WITH 45.7 cm SPACE BETWEEN LARGE AND SMALL PANEL

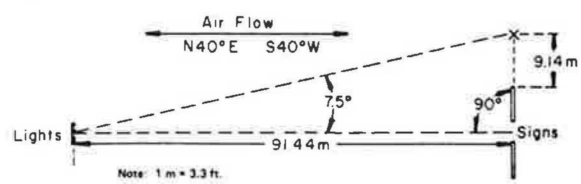
Figure 2. On-site mounting.

were mounted in the field, and the signs were installed at standard mounting height (Figure 2). The headlighting system was installed on a line perpendicular to and centered between the signs so they could be aimed either to a point midway between the signs or to a marker offset 9.144 m (30 ft) from the left-hand sign (Figure 3).

TEST PROCEDURES

Test procedures were varied according to prevalent atmospheric conditions. Normally, observations and photographs were made at 1-h intervals. Some conditions (e.g., a drastic change in the amount of dew)

Figure 3. Layout of test system.



called for the shortening of intervals between observations. The normal period of observations, from 1 h after dark until midnight, was varied to meet conditions and schedules. During early experimental observations in April and May, it was often necessary to wait until after midnight to observe dew or frost conditions. Subsequently, observations were continued only until midnight or until atmospheric and sign conditions stabilized.

Sign positions were interchanged (left to right) in September, approximately at the midpoint of the study, to examine the possibility that the positions of the signs relative to each other were affecting the results.

Sign conditions at the time of each recorded manual observation on each of the 31 nights selected for evaluation were photographed and logged for later corroborative comparison with on-site subjective evaluation of the relative performance of each of the combinations of signing materials. Early experimentation resulted in the selection of photographic techniques that produced color slides of sufficient fidelity to allow experienced viewers to arrive at comparative evaluations nearly identical to those of the on-site observer except in those cases of combinations of signing materials with

Table 1. Total hours of observed dew or frost on test materials.

Item	Hours of Dew or Frost on Material	Percentage of Total Hours of Dew or Frost
Legend		
Direct-applied enclosed lens on enclosed lens, large plywood-backed panel	41	74
Direct-applied enclosed lens on enclosed lens, large aluminum panel	32	58
Direct-applied encapsulated lens on enclosed lens, large plywood-backed panel	24	44
Direct-applied encapsulated lens on enclosed lens, large aluminum panel	18	33
Embossed enclosed lens on enclosed lens, small aluminum panel	35	64
Embossed encapsulated lens on enclosed lens, small aluminum panel	26	47
Embossed encapsulated lens on encapsulated lens, small aluminum panel	20	36
Direct-applied encapsulated lens on encapsulated lens, small aluminum panel	8	14
Button copy on encapsulated lens, large plywood-backed panel	26	47
Button copy on encapsulated lens, large aluminum panel	20	36
Demountable encapsulated lens on encapsulated lens, large plywood-backed panel	18	33
Demountable encapsulated lens on encapsulated lens, large aluminum panel	12	22
Border		
Direct-applied encapsulated lens on enclosed lens, large plywood-backed panel	16	29
Direct-applied enclosed lens on enclosed lens, large plywood-backed panel	24	44
Button copy on encapsulated lens, large plywood-backed panel	49	80
Demountable encapsulated lens on encapsulated lens, large plywood-backed panel	18	33
Embossed enclosed lens on enclosed lens, small aluminum panel	23	42
Embossed encapsulated lens on enclosed lens, small aluminum panel	16	29
Direct-applied encapsulated lens on enclosed lens, large aluminum panel	15	27
Direct-applied enclosed lens on enclosed lens, large aluminum panel	25	45
Button copy on encapsulated lens, large aluminum panel	49	89
Demountable encapsulated lens on encapsulated lens, large aluminum panel	17	31
Demountable encapsulated lens on encapsulated lens, small aluminum panel	16	29
Direct-applied encapsulated lens on encapsulated lens, small aluminum panel	14	25

similar performance. Use of a 2-s exposure of Kodak Ektachrome EH 135 (ASA 160) film and a Bushnell 300-mm lens with blue filter resulted in sign photographic brightness approximating bright headlight illumination at f5.5 and dim headlight illumination at f8.

RESULTS

Total hours of observed dew or frost on each of the combinations of sign materials are given in Table 1. Although 85.5 h of dew or frost occurred at ground level at the test site during the selected periods of observation, only 55 h of dew or frost was observed on one or more of the combinations of sign materials, and a maximum of only 49 h of dew or frost was observed on any single combination. The percentage of total hours of dew or frost on each of the combinations of materials (Table 1) is based on the 55 h of dew or frost formation on the signs. (Temperature, dew point, and humidity data for each of the nights of observation are available on request from the authors.)

One of the most obvious phenomena that recurred throughout the study was the early accumulation of dew on sign materials that had a plywood-backed aluminum panel (see Figure 1, upper right). Dew always formed there first, with greater subsequent total accumulation, and seemed to affect legend performance more seriously. Furthermore, there was a marked difference in subjectively apparent legend performance on the two types of background sheeting material on this sign panel. Legends mounted on the encapsulated-lens reflective materials performed better than legends mounted on the enclosed-lens material.

All combinations of materials appeared to be less affected by frost than by dew. Button copy performed much better under frost conditions than under heavy dew conditions. However, the performance of button copy relative to the performance of the other legend materials degraded with time. It may be that dirt (accretion of atmospheric dust) played a large part not only in the decrease in relative performance of the button-copy materials but also in the noted general degradation in performance of all test materials over time.

A marked difference in the angularity of button and reflective-sheeting materials was noted. The reflective-sheeting legend could be seen easily up to about 30° from center at 91.44 m (300 ft). The button legend showed very little angularity, especially under dew conditions.

At various times, dew- and frost-free areas would appear on the test signs but, unlike the similar bright areas noted on roadside signs, these were almost never in the same places on successive nights. No convincing explanation of this phenomenon or any clear identification of the variables believed to be involved was ever found. Combinations of effects peculiar to the test site and the test installation are believed to have been involved.

Under road conditions, a dew-free area is usually observed on the sign face where the posts are attached to the back of the sign. The posts act as heat sinks. It is possible that air currents and turbulence created by traffic near roadside signs also help to cause the sign face to cool before the posts cool. The experimental signs were generally not exposed to such air turbulence. Furthermore, the size of mounting posts was a factor. For signs of this size, the posts used in practice are usually larger than the ones used in the test installation. The extra heat stored in the larger posts would probably cause this phenomenon to be more stable and pronounced in the case of roadside signs. The noted ephemerally bright (dew- and frost-free) spots were not usually associated with the mounting posts at the test installation.

Forty-five days after the beginning of observations, the relative performances of the various combinations of retroreflective materials were judged to be in the following order (beginning with the last legend to lose reflectability):

1. Button copy (RO) on encapsulated lens, aluminum panel;
2. Button copy (RO) on encapsulated lens, plywood-backed panel;
3. Direct-applied encapsulated lens (FT) on encapsulated lens, aluminum panel;
4. Embossed encapsulated lens (LE) on encapsulated lens, aluminum panel;

5. Demountable encapsulated lens (AD) on encapsulated lens, aluminum panel;
6. Direct-applied encapsulated lens (TH) on enclosed lens, aluminum panel;
7. Demountable encapsulated lens (AD) on encapsulated lens, plywood-backed panel;
8. Embossed encapsulated lens (FT) on enclosed lens, aluminum panel;
9. Embossed enclosed lens (LE) on enclosed lens, aluminum panel;
10. Direct-applied encapsulated lens (TH) on enclosed lens, plywood-backed panel;
11. Direct-applied enclosed lens (RU) on enclosed lens, aluminum panel; and
12. Direct-applied enclosed lens (RU) on enclosed lens, plywood-backed panel.

Final subjective ratings of the legends under light dew conditions after 6.5 months of observation, again in order from best to poorest performance, were as follows:

1. RO on encapsulated lens, aluminum panel;
2. AD on encapsulated lens, aluminum panel;
3. RO on encapsulated lens, plywood-backed panel;
4. FT on encapsulated lens, aluminum panel;
5. LE on encapsulated lens, aluminum panel;
6. TH on encapsulated lens, aluminum panel;
7. AD on enclosed lens, plywood-backed panel;
8. FT on enclosed lens, aluminum panel;
9. LE on enclosed lens, aluminum panel;
10. TH on enclosed lens, plywood-backed panel;
11. RU on enclosed lens, aluminum panel; and
12. RU on enclosed lens, plywood-backed panel.

Two exceptions to these findings were noted during heavier dew formation. The button legend, RO, was about fifth or sixth in order of performance under moderately heavy dew conditions. Under conditions of rapidly forming heavy dew, the button-copy letters exhibited the worst performance (eleventh and twelfth).

Border materials performed very similarly to like legend materials with the notable exception that the button border generally exhibited the worst performance of all the border materials. During the later observations, button-copy borders frequently all but disappeared even under light dew conditions.

Most of the observations under frost conditions occurred during the second half of the study after the sign panels were interchanged (left to right). The superiority of button-copy and encapsulated-lens material was often rather striking. Subjective ratings of legend materials relative to each other under frost conditions, beginning with the last to lose reflectivity, were as follows:

1. RO on encapsulated lens, aluminum panel;
2. RO on encapsulated lens, plywood-backed panel;
3. AD on encapsulated lens, aluminum panel;
4. AD on encapsulated lens, plywood-backed panel;
5. TH on enclosed lens, aluminum panel;
6. FT on enclosed lens, aluminum panel;
7. FT on encapsulated lens, aluminum panel;
8. LE on encapsulated lens, aluminum panel;
9. LE on enclosed lens, aluminum panel;
10. TH on enclosed lens, plywood-backed panel;
11. RU on enclosed lens, aluminum panel; and
12. RU on enclosed lens, plywood-backed panel.

Subjectively rated relative effects of dew and frost on target values of background materials were as follows (the best performance is ranked first):

1. Encapsulated lens on aluminum panel,
2. Encapsulated lens on plywood-backed panel,
3. Enclosed lens on aluminum panel, and
4. Enclosed lens on plywood-backed panel.

CONCLUSIONS

The frequency of some amount of dew or frost formation observed on test sign materials was much greater than expected, i.e., more than 65 nights (9:00 p.m. to midnight) out of the 214 nights (April 18 to November 17) involved. On many selected observation nights when there was noticeable dew on the signs, the formation of ground fog prevented photography essential to the established evaluation procedure. Furthermore, observers frequently could not be present at the test site when atmospheric conditions were suggestive of dew or frost formation. It is conservatively estimated that noticeable dew or frost formation on the test signs, viewed from a distance of 91.44 m (300 ft) under direct headlight illumination, occurred on at least one out of every three nights during the study period. The adverse effects of dew and frost on signs are therefore not frequently imposed on motorists in areas that have climates similar to that of the test site.

Appropriate use of the guidance provided by the following conclusions is thus recommended:

1. Under direct headlight illumination from a distance of 91.44 m and with natural dew formation conditions at the sign face, the performance of encapsulated-lens retroreflective sign materials was found to be far superior to that of enclosed-lens material and equal (for light dew) or superior (for heavy dew) to button copy. Under the same viewing conditions but under natural conditions of frost formation, the performance of button copy was far superior to that of all other test materials; however, viewed from a position offset 1.8 m (6 ft) laterally from the light source, encapsulated-lens material and button copy were almost identical in performance under frost conditions; viewed from greater lateral offset distances, all other test materials were far superior to button copy under both dew and frost conditions.

2. Of the sign materials under study, button-copy legends exhibited the most noticeable continuing degradation in performance under both dew and frost conditions during the 8 months of observations, presumably because of aging (weathering) or accumulation of dirt film or both.

3. The observed decreases in sign legibility and target value because of dew and frost formation were always more pronounced on the plywood-backed sign panel than on the plain aluminum panels.

4. Because of the reduction in dew and frost formation time, encapsulated-lens background material greatly improved the legibility of all legends under all degrees of dew and frost formation observed. This was most noticeable in the case of the directly applied legends: An encapsulated-lens legend on an encapsulated-lens background was far superior to an encapsulated-lens legend on an enclosed-lens background.

5. Subjective ratings of the relative performance of the sign materials under dew and frost conditions were almost imperceptibly affected by aiming the headlights 30° to the left of center of the test sign installation (Figure 3). However, when the headlights were centered on the test sign installation and observers viewed the signs from an angle of 1° to the left or right of center at the light source, both the lack of angularity of button copy and the superiority of the encapsulated-lens materials were quite apparent. Differences in angularity

are of concern in cases such as cab-over-engine trucks that have 254-cm (100-in) driver eye height and misaimed or failed left headlight(s).

6. The performance of all combinations of sign materials appeared to be less affected by frost than by dew.

7. An encapsulated-lens legend on an encapsulated-lens background (the small sign at the bottom right of Figure 1) was less than half as much affected by dew or frost in the case of the directly applied legend (FT) as in the case of the embossed legend (LE). However, this comparative advantage from use of direct applied copy was not evident in the relative performances of direct applied versus embossed borders; direct applied, embossed, and demountable borders exhibited only slight differences in performance, most of which could be explained in terms of other variables such as sign backing, background material, and border material.

8. Reversing the positions of the two sign panel combinations (left to right in Figure 1) had no effect on the subjectively rated relative performances of the signing material combinations.

9. Under the conditions of this study, 80 percent of the noted adverse effects of dew or frost on the conspicuity and specificity of enclosed-lens legends on enclosed-lens backgrounds on plywood-backed sign panels (RU, upper right in Figure 1) could have been avoided through the use of encapsulated-lens legends on encapsulated-lens backgrounds on plain aluminum panels (FT, bottom right in Figure 1).

10. Allowing for normal variations in atmospheric conditions (light dew, rapidly forming heavy dew, and frost) and in signing practice (plywood versus aluminum panels and direct applied versus demountable copy), it is estimated that 50 to 80 percent of the adverse effects of dew and frost could be overcome through the use of encapsulated-lens signing materials.

REFERENCES

1. Manual on Uniform Traffic Control Devices for Streets and Highways. Federal Highway Administration, U.S. Department of Transportation, 1971.
2. F. R. Hanscom. Driver Awareness of Highway Sites With High Skid Potential. Biotechnology, 1974, 204 pp.
3. J. M. Roberts, J. W. Hutchinson, and F. R. Hanscom. Driving Behavior and Expressive Self-Testing Attitudes. Paper presented at 55th Annual Meeting, TRB, 1976.
4. F. R. Hanscom and W. G. Berger. Vehicle Behavior at One Interchange Related to Driver Responses to Guide Signs. Paper presented at 55th Annual Meeting, TRB, 1976.
5. G. J. Alexander and H. Lunenfeld. Some Factors Affecting Reception and Use of Information by Drivers. Public Roads, Vol. 37, No. 1, June 1972, pp. 9-12.
6. G. J. Alexander and H. Lunenfeld. Positive Guidance in Traffic Control. Office of Traffic Operations, Federal Highway Administration, U.S. Department of Transportation, April 1975, 57 pp.
7. Color as Seen and Photographed. Eastman Kodak Co., 1950, 68 pp.
8. W. P. Youngblood and H. L. Woltman. A Brightness Inventory of Contemporary Signing Materials for Guide Signs. HRB, Highway Research Record 377, 1971, pp. 69-91.
9. E. D. Brenning. Evaluation of Reflectorized Traffic Sign Materials Under Controlled Rainfall Conditions. Highway Traffic Safety Center, Univ. of Illinois, Research Rept. 7, Feb. 1973, 73 pp.
10. R. E. Dewar and J. G. Ellis. A Comparison of Three Methods for Evaluating Traffic Signs. Paper presented at 53rd Annual Meeting, HRB, 1974.
11. R. L. Rizenbergs. A Survey on Reflective Sheeting for Highway Signs. Kentucky Department of Transportation, Research Rept. 364, April 1973, 10 pp.
12. R. L. Rizenbergs. High-Intensity Reflective Materials for Signs. Kentucky Department of Transportation, Research Repts. 330 and 368, June 1972 and May 1973.

Publication of this paper sponsored by Committee on Visibility.

Relation Between Sign Luminance and Specific Intensity of Reflective Materials

W. P. Youngblood and H. L. Woltman, 3M Company, St. Paul, Minnesota

Recommendations related to nighttime luminance for traffic signing are not readily translatable from specification or photometric descriptions of the reflective brightness of materials. An investigation of a simple means of translation was undertaken to aid in the proper selection and application of materials where a sign luminance level is desired. The study approach used a photometric determination of specific intensity of the reflective material. The two observation angles common to most highway specifications, 0.2° and 0.5° at -4° entrance angle, were used for determining a broad luminance span for a variety of reflective materials in the common traffic colors. These materials were then installed on a test road where field determinations of sign luminance were also made. The many readings were then correlated by linear regression. These expressions, based on direct observational data, are shown for a variety of shoulder and overhead sign positions, for upper and lower beams, and for the two distances most closely approximating the 0.2° and 0.5° observation angles—183 and 91.5 m (600 and 300 ft). The resulting expressions permit simple computation of either sign luminance or specific intensity for a reflective sheeting.

It is acknowledged that nighttime sign performance is dependent on attention value and legibility. Each factor is related directly to the luminance contrast, the sign with its surround providing attention value and the letters with the sign background providing legibility. Literally, contrast is the luminance difference between an object and its background and is a subjective experience that is given to extreme variation, particularly at night. Excessive stimuli from glare sources, such as opposing headlamps, highway luminaires, and electric advertising, contrast with the generally inadequate luminance needed elsewhere for effective nighttime perception.

In recognition of this, the Manual on Uniform Traffic Control Devices (MUTCD) (1) requires reflectorization or illumination of signs, delineators, and pavement markings. Although the MUTCD requirement has done much to improve visibility, no minimal values are specified and no maintenance of minimal luminance is suggested.

Numerous performance levels of reflective materials are available in various federal and state specifications, and a wide variety of lighting designs and luminaire fixtures exist for compliance with the manual requirement. An obvious difficulty arises in translating reflective material specification values to sign luminances suggested by research for various situations. Although such research has yet to be adopted by the MUTCD, desirable and minimum nighttime levels of sign luminance have been suggested by Lythgoe (2), Smyth (3), Allen and Straub (4), Allen and others (5), Forbes (6), Olson (7), Jainski (8), and other researchers. Such research indicates that increasing sign luminance is required where sign surrounds possess increasing luminance and may vary depending on such factors as the color and size of the sign.

The performance recommendations given in luminance terms are not easily equated to photometric values of reflective material specifications. Reflective luminance has been generally quantified for various materials by Youngblood and Woltman (9). This previous work used a telephotometer at driver eye position and a vehicle of standard dimensions that had carefully aligned headlamps. Careful work from study to study has validated the efficacy of this approach. What was lacking was a complete and relatively direct relation between the variety of photometric test points and sign luminance. Many very pertinent factors are involved in this relation.

Since the efficiency of reflective sheeting varies widely over useful observation (divergence) angles, the resulting relation is expressed as specific intensity (called reflective intensity in certain specifications) and is the luminance in absolute terms versus the observation angle for each type of reflective material under consideration. Observation angle (called divergence angle in certain specifications) is the angle subtended by the headlamps, the sign, and the reflective light beam at the observer. This angle undergoes significant change as the motorist approaches the sign and greatly influences the resulting sign luminance. This angle increases substantially as reading distances for signs shorten. Further, the greater lateral distance of the right headlamp makes the luminance contribution from this source approximately half that of the left lamp at shorter distances. Both changes necessitate separate calculation of the luminance for each headlamp and for each observation angle.

Illuminance depends on the alignment of the sign with the headlamp beam, and its determination requires the location of the reflective device in the appropriate area of the headlamp isocandle diagram for both high and low beams and for typical conditions of highway alignment. Calculation for each lamp is required as is change in sign position or distance. Luminance values are then obtained by application of the inverse square law. Inherent differences in individual lamps are to some extent compensated for by the presence of two or four lamps. However, variation in voltage, lamp misalignment, changes in automobile loading, and specularity of the road surface all contribute to variation in illuminance so that results are not always consistent.

DESIGN OF EXPERIMENT

Most specifications (10, 11) use photometric test points at -4° entrance (called incidence angle in certain specifications), which is essentially perpendicular to the sign surface. The negative angle is for elimination of specular glare in the photometric test, but traffic signing materials today vary little in reflectivity up to angles of ±10°. Observation angles of 0.2° and 0.5° are intended to conform to typical eye headlamp height and sign-reading distances of interest and correspond approximately to distances of 183 and 91.5 m (600 and 300 ft) respectively. These distances were chosen for our observations as most representative of the two observation angles most frequently encountered in specifications.

TEST ROAD

The test road facility is 670 m (2200 ft) long and was designed and constructed to represent a one-way portion of an Interstate roadway. The facility is a straight section with a uniform +0.4 percent grade. The road surface is of comparatively fine-textured asphaltic concrete and is essentially unworn.

POSITION OF SAMPLE PANELS

The sample panels were positioned as shown in Figure 1, the centroids for four positions of signs: for overhead guide signs, 6.55 m (21.5 ft) above the crown of the roadway centered over the right lane; for the shoulder-mounted guide sign, 13.72 m (45 ft) to the right of the lane and 3.05 m (10 ft) above the elevation of the pavement: for the rural shoulder-mounted regulatory warning and advisory signs, 1.83 m (6 ft) above and 3.05 m (10 ft) to the right of the lane edge; for urban shoulder-mounted regulatory warning and advisory signs, 2.44 m (8 ft) above and 0.91 m (3 ft) to the right of the lane edge. These locations represent the center of typical signs and closely correspond to the recommended placement as specified in the MUTCD.

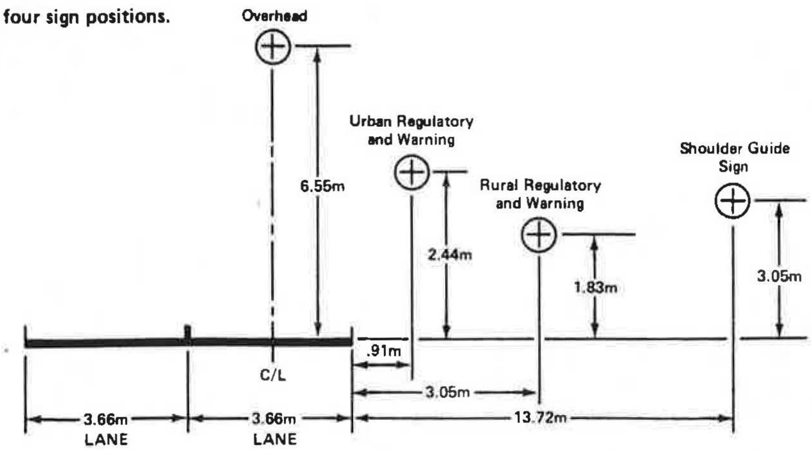
TEST VEHICLE

A full-size station wagon, without tinted glass, was used throughout the test as the primary test vehicle from which measurements were made. Loading conditions of this vehicle were maintained constant throughout the study. The headlamps used were photometrically checked and supplied by General Electric Corporation and conform to the recommended standard for photometrics of the Society of Automotive Engineers (SAE) (12). Two secondary vehicles were also used to check the values obtained with the primary vehicle and to broaden the data base for field luminance. The headlamps of all vehicles were aligned by using the recommended SAE visual aiming method.

SIGNING MATERIALS

The signing materials studied are representative of retroreflective sheeting materials used for traffic-control signs; include silver-white, yellow, orange, red, and green colors; and span a range of specific intensity from 1 to 800 cd/lx/m^2 (1 to 800 cp/fc/ft^2).

Figure 1. Centroids for four sign positions.



The materials used represented three levels of sheeting: enclosed lens, encapsulated lens, and cube corner prismatic. These sheetings, in each color, were further attenuated with up to 10 thicknesses of clear transparent acrylic plastic for purposes of broadening the luminance range from 3 to 15 determinations/color by adding successive thicknesses of plastic after unobstructed readings had been made.

PHOTOMETRIC INSTRUMENTATION

Determinations of specific intensity per unit area of materials used in the outdoor test, including transparent overlays, were made in a 15-m (50-ft) laboratory darkroom equipped for routine photometric testing. The equipment and the procedure used conform to Federal Test Method Standard 370 (13).

Determinations of specific intensity per unit area were made at observation angles of 0.2° and 0.5° at -4° entrance, corresponding to two reflective intensity specification values most representative of the sign-luminance determinations. The photometric equipment uses a 2856 K source and has a photocell corrected for linearity of response. In the National Bureau of Standards (NBS) collaborative tests of reflective materials (14), this equipment has proved to be very close to the median of values reported by all laboratories in the NBS program.

Luminance measurements were made with a Gamma Scientific model 2000 telephotometer. This instrument, which has a transistorized photomultiplier and electrometer amplifier, independent battery power supply, five acceptance angles, a measurement span from 0.003 to 100 000 cd/m² (0.01 to 30 000 ft·L), photopic color correction, and internal standardization and calibration, is suited for such measurements. At the outset and at the conclusion of the tests, the instrument was calibrated with an NBS standard source and over a number of tests averaged ±2.5 percent.

DISCUSSION OF RESULTS

Plotting of sign-luminance measurements versus specific-intensity data reveals a linear relation that differs slightly depending on the chosen shoulder position of the sign and varies quite significantly depending on the beam mode used or if the overhead sign position is used. In the testing, the color of the reflective material was not an apparent variable except that color results in a differing value of specific intensity.

Computer analysis by use of a least squares regression was performed to determine both the linear and exponential fit for a given set of data points. Forty-eight or more pairs of data were analyzed for each linear expression. In each of the sign-luminance specific-intensity determinations given in Table 1, the following expression is used:

$$y = ax + b \quad (1)$$

or

$$x = (y - b)/a \quad (2)$$

where

- y = sign luminance (cd/m²);
- a = slope of the line;
- x = specific intensity of the reflective material (cd/lx/m²); and
- b = constant.

r^2 = quality of fit with the data; it is that portion of the variability in the data that is explained by the regression equation.

As an example, sign luminance is desired for a yellow warning sign in the urban shoulder location when viewed from 91.5 m (300 ft) on low beams. When measured at 0.5° observation and -4° entrance in the laboratory, a material has a specific intensity of 110 cd/lx/m² (110 cp/fc/ft²). From Table 1, the appropriate formula is $y = 0.13x - 0.45$; thus, $y = 0.13 \times 110 - 0.45 = 13.85$ cd/m² (4.0 ft·L) sign luminance.

The 183-m (600-ft) distance is only related to the 0.2° observation angle, and the 91.5-m (300-ft) distance is related to the 0.5° observation angle. These relations are appropriate and must be kept in mind in attempting to predict sign performance.

It should be pointed out that the above relations are appropriate for the typical domestic automobile and headlight and should not be translated to vehicles that have widely differing headlamps or headlamp-to-eye-height distances. The relations hold for colors tested by the authors and dirty and weathered signs but not dirty headlamps or windshields. Dirty or weathered signs must be evaluated with a portable photometer such as a Gamma model 910 or be photometrically evaluated in the laboratory.

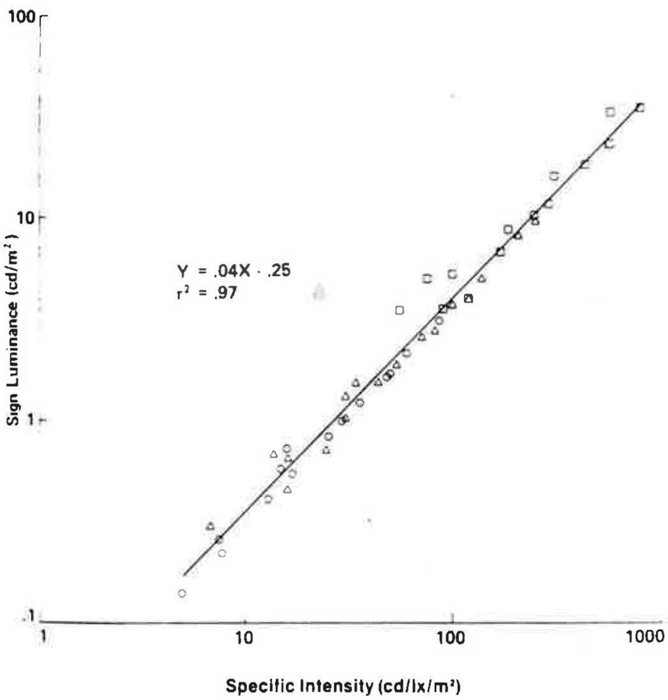
By substituting values and solving for x, the specific intensity of the sheeting can be determined if a predetermined sign luminance is desired. This procedure can aid in the selection of the appropriate reflective material for the sign application. The typical data points that represent sign luminance versus specific intensity of the reflective material for one set of viewings are

Table 1. Equations for sign luminance and specific intensity for various sign locations and beam modes.

Sign Position	Distance (m)	Headlamp Beam	Specific Intensity ($cd/lx/m^2$)	Sign Luminance (cd/m^2)	r^2
Urban shoulder	183	Upper	x = 0.86y + 2.99	y = 1.14x - 0.64	0.98
		Lower	x = 23.16y + 9.75	y = 0.04x - 0.25	0.97
	91.5	Upper	x = 0.24y + 7.62	y = 3.92x - 20	0.94
		Lower	x = 7.47y + 4.07	y = 0.13x - 0.45	0.98
Rural shoulder	183	Upper	x = 1.01y - 0.28	y = 0.96x + 3.72	0.97
		Lower	x = 20.42y + 8.22	y = 0.047x - 0.20	0.97
	91.5	Upper	x = 0.25y + 7.43	y = 3.82x - 21	0.94
		Lower	x = 5.84y + 4.03	y = 0.16x - 0.41	0.96
Shoulder guide	183	Upper	x = 3.07y + 4.80	y = 0.32x + 0.36	0.97
		Lower	x = 3.78y - 4.48	y = 0.25x + 3.42	0.96
	91.5	Upper	x = 2.11y + 13.60	y = 0.47x + 6.02	0.99
		Lower	x = 8.84y + 11.86	y = 0.11x - 1.23	0.99
Overhead	183	Upper	x = 3.45y - 9.71	y = 0.29x + 3.66	0.98
		Lower	x = 52.04y - 12.38	y = 0.02x + 0.29	0.98
	91.5	Upper	x = y - 12.65	y = 1.00x + 12.65	0.98
		Lower	x = 33.33y + 1.67	y = 0.03x - 0.05	0.99

Note: 1 m = 3.3 ft, 1 $cd/lx/m^2$ = 1 $cp/fc/ft^2$; and 1 cd/m^2 = 0.29 ft L.

Figure 2. Sign luminance versus specific intensity for urban sign location at sight distance of 183 m (600 ft) using lower beam headlamps.



shown in Figure 2 together with the linear equation that has the calculated best fit.

CONCLUSIONS

To aid in translating from photometric determinations of specific intensity per unit area of reflective sheetings to the reflective performance of the sign in place, the study examined the relations in a field-laboratory series of tests. Determinations of nighttime sign luminance were made from the driver's eye position in a standard-sized passenger automobile with carefully selected normal headlamps. Measurements were made on a smooth tangent roadway by using a laboratory telephotometer at distances of 183 and 91.5 m (600 and 300 ft). Reflective samples were mounted in typical sign positions. The reflective materials used represented specific intensities from 1 to 800 $cd/lx/m^2$ (1 to 800 $cp/fc/ft^2$) in silver-white, yellow, orange, red, and green.

Specific intensities per unit area were determined for the same materials by standard laboratory photometric methods. Determinations at the observation angle of 0.2° were correlated with 183-m (600-ft) luminance readings and those at a 0.5° observation angle with 91.5-m (300-ft) luminance readings. A linear regression equation was determined for each viewing condition. The resulting equations established the relation between sign luminance and the specific intensity of reflective materials for each distance, sign, and headlamp position.

Should minimum sign luminances be established, or if the research cited previously is used to establish desirable sign luminances, ready translation from photometric values to sign luminance is available in convenient form.

REFERENCES

1. Manual on Uniform Traffic Control Devices for Streets and Highways. Federal Highway Administration, U.S. Department of Transportation, 1971.
2. R. J. Lythgoe. Medical Research Council Report 173. London, England, 1929.
3. J. S. Smyth. The Brightness and Legibility at Night of Road Traffic Signs. Trans., Illuminating Engineering Society, Vol. 12, No. 4, 1947.
4. T. M. Allen and A. L. Straub. Sign Brightness and Legibility. HRB, Bulletin 127, 1955, pp. 1-14.
5. T. M. Allen, F. N. Dyer, G. M. Smith, and M. H. Janson. Luminance Requirements for Illuminated Signs. HRB, Highway Research Record 179, 1967, pp. 16-37.
6. T. W. Forbes. Luminance and Contrast Requirements for Legibility and Visibility of Highway Signs. Michigan State Univ., East Lansing, Pt. 3, Final Rept., April 1975.
7. P. L. Olson and A. Bernstein. Determine the Luminous Requirements of Retroreflective Highway Signing. Highway Safety Research Institute, Univ. of Michigan, Ann Arbor, NCHRP Project 3-24, 1977.
8. I. P. Jainski. The Influence of Surface Color on the Perception, Recognition and Reading of Non-Self-Luminous Traffic Signs. Strassenverkehrstechnik, West Germany, Vol. 21, No. 2, 1977.
9. W. P. Youngblood and H. L. Woltman. A Brightness Inventory of Contemporary Signing Materials for Guide Signs. HRB, Highway Research Record 377, 1971, pp. 69-91.
10. Standard Specifications for Construction of Roads and Bridges on Federal Highway Projects. Federal Highway Administration, U.S. Department of Transportation, FP-74, 1974.
11. Sheeting and Tape, Reflective: Nonexposed Lens, Adhesive Backing. Federal Specification LS-300B, July 1974.

12. Lighting Equipment and Photometric Tests. SAE, Handbook Supplement HS34, March 1974.
13. Instrumental Photometric Measurement of Retroreflective Materials and Retroreflective Devices. Federal Test Method Standard 370, March 1, 1977.
14. Collaborative Reference Program—Color and Appearance. National Bureau of Standards, U.S. Department of Commerce, Retroreflectance Rept. 3, Feb. 1978.

Publication of this paper sponsored by Committee on Visibility.

Evaluation of Daytime High-Visibility Aids for Motorcyclists

Norman Ashford, P. Stroud, C. Kirkby, and N. S. Kirk, Loughborough University, England

The results of a survey of consumer attitudes toward such conspicuity aids for motorcyclists as jackets, waistcoats, sleeves, and slipovers are reported, and the results of laboratory and field trials conducted to determine the effectiveness of such conspicuity aids in facilitating the detection of motorcyclists are reported. These results are based on the first three years of a continuing research project. The user attitude survey indicates serious design problems with some types of conspicuity aids and, for most materials, a severe lack of fastness of both color and fluorescence. The laboratory trials indicated an inverse logarithmic relation between the projected area of fluorescent color and mean detection time.

To examine some of the problems associated with the design, use, and effectiveness of high-visibility aids and clothing for daytime use by motorcyclists, the U.K. Transport and Road Research Laboratory has sponsored a 3-year evaluation program that has been carried out by the Institute for Consumer Ergonomics and the Department of Transport Technology at Loughborough University. This paper briefly discusses the three principal research areas investigated in this project:

1. An evaluation of user attitudes to the types of clothing and other conspicuity aids currently in production and the subsequent design of more suitable clothing (1),
2. A laboratory simulation of the effectiveness of high-visibility aids in the daytime detection of motorcyclists (1), and
3. Field trials to determine the effect of such high-visibility aids on gap acceptance by drivers (2).

These research areas carried out over a period of 3 years form three parts of a continuing program of research into the conspicuity of two-wheeled vehicles that in the long term will embrace both motorized and nonmotorized vehicles under both daytime and nighttime conditions.

USER SURVEY

Study Design

There is strong evidence that, although motorcyclists can make themselves more visible by wearing such fluorescent clothing as slipovers, waistcoats, or jackets, there is some consumer reticence toward using these conspicuity aids. Generally the number of riders wearing fluorescent clothing is very small; in an observational survey carried out in conjunction with this work, only 1.5 percent of the sample (N = 2842) were observed to be wearing any type of high-visibility clothing. To examine this problem in greater depth, a series of discussions on attitudes was carried out with groups of motorcyclists. This was followed by a survey of users' opinions on safety clothing. The survey attempted to

1. Establish the perceived effectiveness of different safety clothing,
2. Isolate particular problems of use,
3. Evaluate the acceptability of high-visibility clothing,
4. Determine users' willingness to purchase such garments, and
5. Evaluate the fastness of the fluorescence and color of the clothing.

A number of different styles of safety-related clothing were purchased and distributed free of charge to motorcyclists in four different areas in the United Kingdom. After three months of use, the motorcyclists were requested to complete an evaluation questionnaire. A large range of safety clothing was obtained, and from this range 19 items were selected for evaluation on the basis of the following five criteria:

1. Style—slipover, waistcoats, jackets, and sleeves;
2. Method of fastening—zip, Velcro, ties, buttons, elasticated sides, and press studs;
3. Material—Wavelock PVC, PVC-coated woven fabric, Webb-lite;
4. Color—red-orange to orange range plus Saturn yellow; and
5. Cost.

Altogether, 924 items of clothing were distributed in five population centers: Swindon (290), Peterborough (88), Nottingham (150), Manchester (113), and Loughborough (283). As the clothing was distributed, anthropometric measurements were taken from the users. Because sleeves were an unpopular option, 32 pairs of sleeves were given to respondents who were also given a slipover or a waistcoat. Therefore, only 892 volunteers received the 924 items. Three months after the date of distribution, the volunteers were each sent a copy of the evaluation questionnaire. Three reminders were sent to nonrespondents. The response rate obtained from the participants was excellent; 93 percent replied to the questions in the areas indicated in the table below:

Area	Question
Details of machine	Type of machine
	Engine size
	Accessories fitted
Use of machine	Frequency of use for different activities
	Use for longer journeys
	Annual distance
Views on safety clothing tested	Frequency of use for different activities
	Whether or not clothing still being worn
	Reasons for no longer wearing clothing
	Perceived effectiveness
	Clothing worn under safety clothing
	Storage of safety clothing
	Type of fastening
	Ease of doing up and undoing
	Fastening damage
	Ease of putting on and taking off
	Effect of cold weather
	Adequacy of adjustment
	Satisfaction with length of safety clothing
	Maximum speed at which safety clothing worn
	Inconvenience caused at that speed
	Effect of wind stress
	Interference of safety clothing with riding
	Need for cleaning of safety clothing
	Frequency and ease of cleaning
	Suitability for use throughout the year
	Embarrassment caused by wearing safety clothing
	Previous use of high-visibility clothing
	Value of safety clothing
	Value of other types of safety clothing
	Reference for different types of safety clothing
	Willingness to purchase types again
	General comments
Recent accident experience	Incidence of recent multivehicle collisions
	Frequency
	Use of safety clothing at time of accident

Results

A comparison was made between the distribution of motorcycle ownership by engine capacity for the study volunteers versus the known pattern of ownership for the general licensed population. The survey population was found to be underrepresentative of riders of small machines and overrepresentative of riders of large machines. It was felt that this was attributable to a high incidence of "enthusiasts" among the volunteers; it was not, however, considered to be an invalidating bias. The average distance traveled was approximately 5790 km (3600 miles), which indicated a normal level of use among the respondents.

After the 3-month trial period, 75.5 percent of the respondents indicated that they were still wearing their test clothing. Of the remainder, who no longer used the clothing, it was found that 50.3 percent had stopped wearing it in the first month of the trial. A variety of reasons were given for discontinuance: too troublesome and inconvenient (27.0 percent), no longer had a motorcycle (19.5 percent), considered it was for nighttime use only (8.8 percent), had purchased another item of safety clothing (6.5 percent), embarrassment (5.1 percent), and illness (1.4 percent).

The survey indicated that the overwhelming majority of motorcyclists wore either motorcycle jackets (45.5 percent), anoraks (24.1 percent), or three-quarter-length coats (18.9 percent) under the safety clothing. The incidence of motorcycle jackets was much higher than it had been in a complementary observational study conducted throughout the United Kingdom in which only 27.6 percent of all riders were seen to be wearing motorcycle jackets. There is a strong positive correlation between the size of the rider's machine, expressed in terms of engine capacity, and the wearing of a motorcycle jacket. The high incidence of jackets in this study is mainly accounted for by the bias toward large machines.

A large proportion of those surveyed (81.9 percent) considered the clothing to be suitable for use throughout the year. Among the remainder, 42.4 percent considered that the clothing would cause sweating in summer, 26.5 percent did not feel it was necessary in the long daylight hours of summer, and 9.9 percent indicated that the clothing was too large to be used comfortably over summer clothing.

Failure to wear safety clothing is frequently imputed to the embarrassment caused by its color, material, and styling. Even among those who volunteered to participate in this work, 25.3 percent admitted to embarrassment. This was not sensitive to particular options. A number of reasons were given for embarrassment: initial self-consciousness or embarrassment caused by others' comments when the clothing was first worn (33.3 percent), a feeling that one was in a minority and consequently too conspicuous (17.4 percent), disquiet over the style of the clothing (10.9 percent), admission to particular embarrassment when the rider wearing the clothing was not riding the motorcycle (22.9 percent), or a feeling that fluorescent clothing was unnecessary in daylight conditions (4.5 percent).

It was found that 18.6 percent of respondents had worn this type of clothing before. This is very much higher than the 1.5 percent who were observed to be wearing such clothing in the complementary study and reflects the level of interest and enthusiasm of those who chose to participate.

Table 1 summarizes general comments about the 19 options. Table 2 gives a summary of users' evaluations of the options and converts their comments into ratings.

The behavior of the fluorescent materials under prolonged exposure to light was tested for each of the 19 options by exposing five 7-cm squares taken from each garment. One set was designated "control," and the other patches were attached to a frame and exposed horizontally on a flat roof. The control samples were measured for International Commission on Illumination (CIE) chromaticity and luminance values. The illuminant approximated the D_{65} light source, and measurements were taken on one thickness of material backed by a standard grey tile that had a luminance of 0.59. After 3 months and 6 months of exposure, further sets were sent for measurement. The patches were washed monthly and immediately before measurement. Those exposed for 9 months were not measured since all colors had faded and in some cases the fabric had disintegrated. The control set, having been kept in a light-proof place, was remeasured; it was found that there was no change in chromaticity coordinates in these control pieces.

Table 3 gives the readings for the three sets: control, 3 months of exposure, and 6 months of exposure. The very large changes in color are shown in Figure 1, in which a selection of large shifts in CIE chromaticity coordinates, shown approximately in the center of the chart, indicate a desaturation of color. (Only the measurements for the control and 6-month samples are shown in the figure. These are joined by straight lines only for clarity and not to represent the locus of fading. All samples desaturated during exposure, and their plotted points moved toward the measuring illuminant, D_{65}.) PVC-coated materials performed relatively better, and option 8 performed best. The fading of nylon patches was rapid: After only 3 months they were almost transparent.

Table 1. General comments on 19 safety options from rider survey.

Number	Option	Comments
1	40-cm sleeves, PVC-coated woven fabric	Sleeves not a popular option, mainly wanted by those who wished to be more conspicuous when indicating turns; not easy to put on, especially when stiffened by cold weather; complaints of reduced circulation because of tightness around wrists and elbows
2	Slipover, PVC-coated woven fabric, lace, tie, and elasticated loop fastening	Main shortcoming a difficulty in fastening with elasticized loop; frequent breaking of stitching of fastening to fabric; although seldom used at high speeds, ballooning and flapping gave large problems
3	Slipover, embossed PVC, unattached lace ties through eyelets	Many complaints about short, easily lost laces, which were also difficult to do up in cold weather; flapping at speed; materials ripped easily around eyelets; head opening too small to pass helmet; option rode up motorcyclist's back
4	Slipover, Wavelock PVC, plastic strip fastening with buttons and buttonholes	Worst fastening failure of any option (50 percent in 3 months); buttonholes main failure but also strap and button failure; fastening task difficult in cold weather, especially with gloves; material curled; adjustment provided considered fairly good
5	Slipover, Webb-lite fastening by stitched lace ties	Ties again caused many complaints; knots became tight; difficult to undo with cold or gloved hands, especially when wet; difficult to clean, subject to billowing, and frequently considered too short
6	Slipover, acrylic nylon, fixed elastic sides	Difficulty with putting garment on with fixed elastic sides; damage to fastening frequently caused by strain of putting on and taking off; garment billowed and rode up; head opening too small to accommodate helmet
7	Slipover, embossed PVC, fastening with buckles and canvas straps, canvas shoulder straps	Tearing around stitching of straps to PVC buckle; fastening difficult, especially in cold weather; insufficient adjustment in canvas straps when worn over winter clothing
8	Slipover, PVC-coated woven fabric, press-stud fastening on elastic strip	Press-stud fastening easier to do up than many other types; head opening too small for helmet; longer back portion flapped when riding and doubled over
9	Slipover, Wavelock PVC, Velcro flap fastenings	Most satisfactory response of any slipover; easy tab fastening; billowing and flapping might have been more frequent with greater exposure to high speeds; head opening too small for helmet
10	Slipover, embossed PVC, small Velcro tab fastening	Fastening more difficult to use than those of option 9; damage at fixing of fastening to PVC; many complaints about billowing and flapping
11	Short waistcoat, PVC-coated woven fabric, Saturn yellow, Velcro flap fastening	68 percent of wearers complained of shortness; tight and uncomfortable over winter clothing; frontal high-visibility areas considered insufficient; equal number of comments for and against color
12	Waistcoat, PVC-coated woven fabric, front and side fastening by press stud, open sides	Fastening not difficult; subject to billowing and flapping; front area obscured by flapping up
13	Waistcoat, Wavelock PVC, front fastening by press studs	Generally well received; small press studs difficult in cold weather; without adjustment, could be tight over winter clothing
14	Waistcoat, woven nylon fabric, zip fastening	Most satisfactory of waistcoat options; easy and convenient to use, lightweight, easily stored; difficult to clean; zip tab difficult to grip
15	Waistcoat, Webb-lite, Velcro strip fastening, Saturn yellow	Velcro strip poorly attached, easily damaged, required difficult alignment; difficult to clean and heavy and difficult to store; material holds water; tight and nonadjustable over winter clothing
16	Waistcoat, PVC-coated woven fabric, fastening by four large plastic buttons	Stiff and not easy to store; fastening difficult in cold weather as material stiffens; considerable fastening damage observed with use
17	Overjacket, woven nylon, zip fastening, elasticized cuffs and waist	60 percent of wearers found option too short; zip tab found fiddly; difficult to accommodate bulky clothing; pocket found very desirable
18	Three-quarter-length coat, PVC-coated woven fabric, press-stud fastening	Difficult to put on over motorcycle clothing; fastening fiddly; ballooning and violent collar flapping at higher speeds found very disconcerting
19	Hooded Anorak, acrylic-proofed nylon, fastening by double-ended zip, inner storm cuffs fastened of Velcro, drawstrings around hood and bottom of garment	Generally highly acceptable and worn by many when not riding motorcycle; at speed, hood flapped violently; hood considered unnecessary by many; double-ended zip difficult to fasten

Table 2. User ratings, chromaticity coordinates, and unit cost of 19 options.

		Questionnaire Rating						CIE Tristimulus Coordinates[a]			
Option	Actual Length (cm)	Satisfaction With Length	Speed Exposure	Inconvenience at Maximum Speed	Ripping Caused by Wind Stress	Interference With Riding	Ease of Cleaning	x	y	Y (%)	Unit Cost (£)[b]
1	38	Very good	Less than adequate	Fair	Very good	Less than adequate	Less than adequate	0.598	0.337	60.2	0.73[c]
2	54.5	Fair	Poor	Poor	Good	Good	Fair	0.590	0.336	64.4	0.62[c]
3	48	Less than adequate	Less than adequate	Fair	Good	Fair	Fair	0.610	0.338	59.1	0.81[c]
4	61	Very good	Less than adequate	Fair	Poor	Good	Good	0.556	0.367	75.1	0.53[c]
5	46	Poor	Less than adequate	Less than adequate	Good	Good	Poor	0.578	0.353	68.9	1.38[c]
6	48	Poor	Less than adequate	Fair	Very good	Very good	Good	0.599	0.342	51.7	1.08[d]
7	48	Less than adequate	Fair	Good	Fair	Good	Good	0.592	0.357	66.4	1.62[c]
8	51 (front), 66 (back)	Very good	Less than adequate	Fair	Very good	Good	Very good	0.646	0.339	45.8	1.35[d]
9	66	Good	Less than adequate	Good	Very good	Fair	Very good	0.595	0.365	72.1	2.00[c]
10	58	Fair	Less than adequate	Fair	Good	Good	Fair	0.601	0.365	70.9	1.16[c]
11	41	Poor	Very good	Good	Very good	Fair	Fair	0.402	0.552	116.0	1.15[c]
12	59	Fair	Less than adequate	Poor	Very good	Very good	Fair	0.639	0.329	41.1	1.41[c]
13	58	Fair	Good	Good	Good	Good	Fair	0.597	0.368	67.2	0.84[c]
14	61	Good	Fair	Very good	Very good	Very good	Less than adequate	0.613	0.333	46.2	2.32[c]
15	68.5	Very good	Good	Fair	Very good	Fair	Poor	0.385	0.520	93.3	4.09[c]
16	68.5 (small), 70 (medium), 71 (large)	Good	Good	Very good	Very good	Fair	Good	0.614	0.335	37.7	2.02[c]
17	53 (front), 58 (back)	Poor	Very good	Fair	Very good	Good	Less than adequate	0.584	0.375	52.7	3.25[c]
18	85	Very good	Fair	Poor	Very good	Poor	Fair	0.596	0.332	63.3	3.96[c]
19	76	Very good	Fair	Poor	Good	Good	Fair	0.611	0.330	43.7	3.90[c]

Note: 1 cm = 0.39 in.

[a]Coordinates of International Commission on Illumination. [b]Approximate 1977 exchange rate of 1 £ = U.S. $1.80. [c]Wholesale. [d]Retail.

LABORATORY SIMULATION OF EFFECTIVENESS OF HIGH-VISIBILITY AIDS

One of the complementary studies to the survey of rider attitudes toward conspicuity aids was a controlled laboratory examination of the effectiveness of different aids. After a literature survey and extensive discussions with others working in the areas of conspicuity and visibility, it was decided that the most suitable laboratory technique was likely to be the tachistoscope.

The three principal factors that affect target recognition were considered to be the target itself, the background, and the method of presenting a stimulus. A number of methods of presenting the target were considered, tried, and eventually rejected. Among those rejected were (a) colored target stimuli on plain backgrounds, (b) colored target discs on photomontages of street scenes, (c) abstract backgrounds with targets added, and (d) artist's impressions of typical street scenes with superimposed figures of different sizes. The first two techniques were tested and abandoned because of the ease of target detection and the failure of the techniques to discriminate between target options; the latter two techniques were abandoned because of lack of realism.

Another problem that affected the first approaches to the laboratory work was the presence of fluorescent stimuli. When introduced into the tachistoscope, a small fluorescent patch of color did not give its true fluorescent color in the absence of the ultraviolet radiation of normal daylight. In the final test procedure, this problem was avoided by testing options of identical color in the tachistoscope so that the differences in detection times would not include the color effect.

Experimental Stimuli Material

The options tested included clothing and machine-based items—namely, leg shields, headlamp covers, sleeves, waistcoats, jackets, and helmets. The control option was a motorcyclist wearing a black open-face helmet, a dark green Belstaff motorcycle jacket, black gloves, and blue denim jeans. Motorcycle and rider were photographed in nine urban road sites selected to give a range of backgrounds and traffic densities.

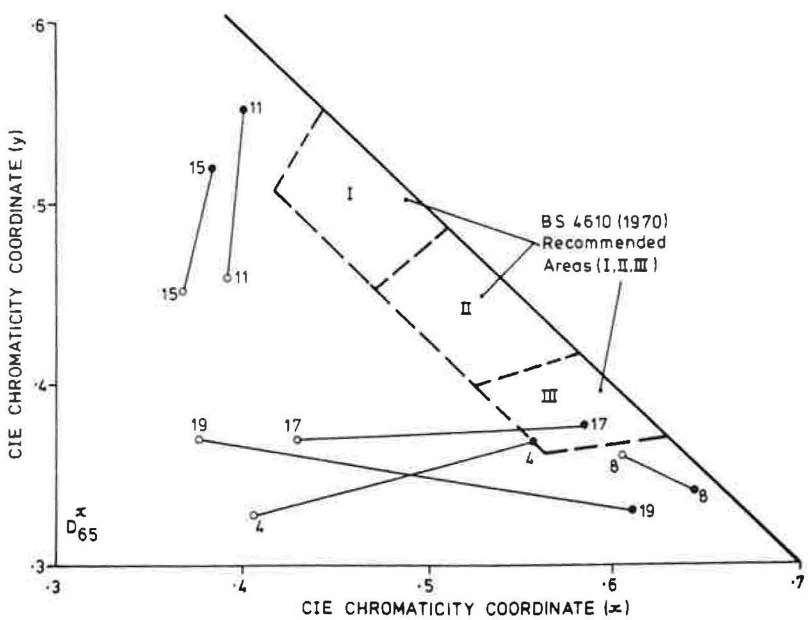
Apparatus

Figure 2 shows the layout for the apparatus used in the laboratory trials. The technique used the back projec-

Table 3. Color changes on exposure for three sets of fluorescent materials.

	Control			3 Months of Exposure			6 Months of Exposure		
Option	x	y	Y (%)	x	y	Y (%)	x	y	Y (%)
1	0.5980	0.3320	60.15	0.5842	0.3316	49.05	0.5402	0.3423	50.20
2	0.5900	0.3361	64.45	0.5759	0.3334	54.10	0.5141	0.3499	61.20
3	0.6095	0.3380	59.15	0.5924	0.3348	53.10	0.5065	0.3178	53.30
4	0.5560	0.3679	75.15	0.5261	0.3759	64.40	0.4051	0.3278	62.35
5	0.5785	0.3559	68.95	0.5640	0.3568	56.50	0.5221	0.3666	61.60
6	0.5991	0.3421	51.70	0.5548	0.3574	36.05	0.4189	0.3676	42.60
7	0.5923	0.3572	66.35	0.5675	0.3703	56.40	0.4219	0.3802	68.00
8	0.6457	0.3392	45.75	0.6373	0.3433	37.80	0.6057	0.3598	38.10
9	0.5953	0.3659	72.05	0.5550	0.3727	61.10	0.4328	0.3813	57.25
10	0.6011	0.3655	70.90	0.5673	0.3778	62.60	0.4074	0.3858	72.65
11	0.4022	0.5521	115.50	0.4155	0.5310	92.15	0.3937	0.4606	73.80
12	0.6385	0.3292	41.15	0.6210	0.3325	39.30	0.5606	0.3466	49.15
13	0.5965	0.3686	67.25	0.5754	0.3743	60.75	0.4948	0.3952	53.20
14	0.6128	0.3336	46.25	0.5582	0.3481	33.70	0.3740	0.3650	45.90
15	0.3850	0.5188	93.25	0.3869	0.4880	70.35	0.3699	0.4505	69.10
16	0.6136	0.3349	37.65	0.6044	0.3384	40.45	0.5165	0.3577	47.15
17	0.5844	0.3755	52.75	0.5510	0.3770	34.70	0.4288	0.3696	41.10
18	0.5959	0.3329	63.35	0.5841	0.3324	54.10	0.5278	0.3465	57.50
19	0.6110	0.3306	43.75	0.5531	0.3478	31.85	0.3791	0.3696	44.00

Figure 1. Section of CIE chromaticity chart showing readings for control materials and materials exposed for 6 months.



tion of slides. Because it was recognized that the color rendition of film is not perfect, single-color targets were used to eliminate any color effect. Color-reversal film (35-mm) was used and presented by means of a tachistoscopic slide projector controlled by the subject that back-projected the image onto a screen in front of the subject for as long as the subject's control button was depressed. This time of presentation was recorded on an electronic digital timer. The experimenter was able to advance the slide magazine by use of the slide control button. The subject was seated 1 m (3.3 ft) from the screen, and the visual angle of the motorcycle image approximated a real-world viewing distance of 92 m (300 ft).

Figure 2. Layout of apparatus used in laboratory tests.

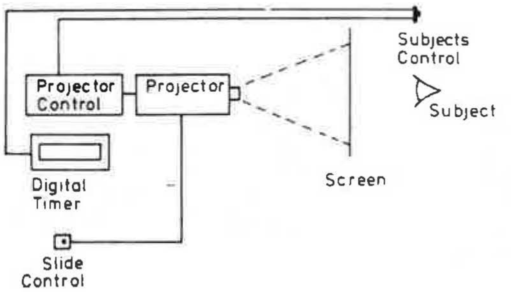


Figure 3. Mean detection and recognition times in pilot experiments.

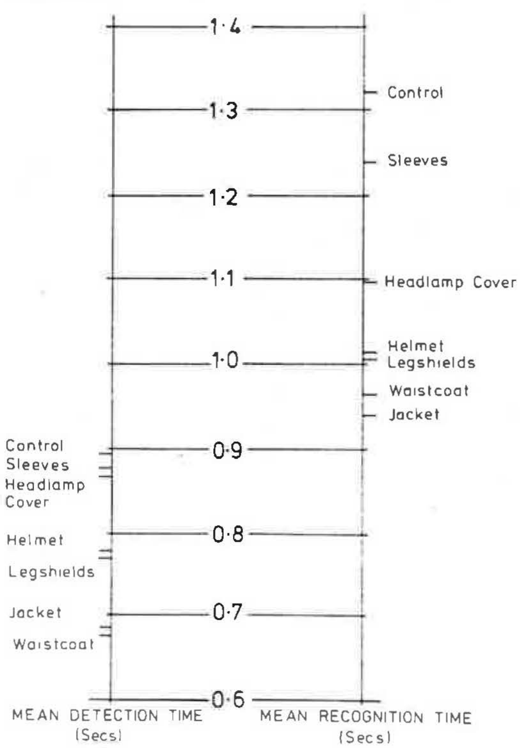
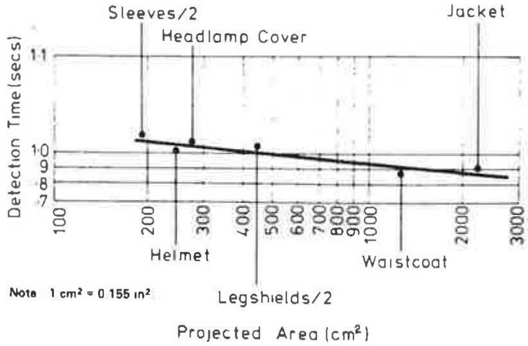


Figure 4. Relation between projected fluorescent area and detection time (log-log scale).



Pilot Trials

Pilot experiments were carried out to validate the apparatus and to determine the best form of task to be given to the subjects. The two tasks set were detection and recognition. Subjects given the detection task were instructed to press their control button to project the traffic scene and then to release it on detecting a motorcyclist. If they could not see a motorcyclist in the scene, they were to release the button and inform the experimenter. Subjects given the recognition task were instructed in such a way that they could identify the control and the six high-visibility options and identify them by their associated names. The recognition task was similar to the detection task except that, after observing the motorcyclist in the scene and releasing the hand button, the subjects were required to state which option was shown in the photograph.

The 45 subjects (24 male and 21 female) were mainly students and university staff. All were given an Ishihara color vision test, and no defects were recorded.

The ordering and grouping of the mean recorded times for the two tasks are shown in Figure 3. The results are clearly of similar form although the range and value of the times vary. From these results, it was decided that the method of presentation permitted discrimination and presented options in an order that might be expected, i.e., in which the large areas of the waistcoat and jacket were perceived quickest and options with smaller areas took longer.

Of the two tasks used, the recognition task presented the most problems. Because many subjects did not release their hand button immediately on realizing which option was presented, recorded recognition times appeared to be artificially increased and unrepresentative of the true time. Frequently, the scene was retained while the subject checked with a reference set of photographs for the correct name of an option.

The validity of the results of the detection tests was confirmed by the inclusion of scenes that contained no motorcyclist. It was found that subjects given the detection task correctly reported no motorcycle present for all 10 blank slides presented. From this, it was concluded that the detection times were valid and not those times produced by subjects who released the hand button after a short period without actually perceiving the motorcyclist.

The form of the results of the pilot trials indicated that the experimental technique could be adopted for the main laboratory trials and that the detection task was most suitable for determining the conspicuity of motorcyclists.

Main Laboratory Trials

The main laboratory trials conducted to measure the relative effectiveness of high-visibility fluorescent orange were conducted on the same equipment and with the same form of stimuli as those used in the pilot trials. The 72 experimental slides were presented in random order, and half of the slides were reversed to ensure balanced presentation on both the left and right sides of the screen.

The technique developed for the laboratory trials had proved to be satisfactory with respect to the ease with which experimental stimuli could be presented to the subjects. It was found that the method could be easily replicated and that new options for evaluation could be added for direct comparison with options already tested. It is important to emphasize that, because of limitations in the photographic reproduction of color, the technique can be used only to compare different options of the same color. The inability of film stock to reproduce fluorescent colors is particularly critical.

The mean detection times for the seven options across all nine sites are as follows: control, 1.090 s; sleeves, 1.116 s; leg shields, 1.048 s; jacket, 0.896 s; headlamp cover, 1.070 s; helmet, 1.008 s; and waistcoat, 0.880 s. Dunnett's statistic at the 0.05 level indicated that the jacket and waistcoat produced detection times faster than the control whereas all other options did not. A less stringent test using the t-statistic for individual comparisons showed that the jacket, waistcoat, and helmet produced times faster than the control. Although the jacket had an area almost twice that of the waistcoast—2260 versus 1270 cm^2 (350 versus 197 in^2)—no significant difference could be found in their mean detection times. This result is likely to arise from two effects that act either separately or in concert:

1. There is a cut-off point in the detection time-area relation beyond which an increase in the size of the area does not result in a decrease in detection time.
2. The smaller area of the waistcoat is compensated for by a contrast with the areas of the arms and shoulders in dark motorcycle clothing. Contrast in this case, and consequently visibility, are therefore not so dependent on background as they are for the jacket option.

These trials indicated the following inverse logarithmic relation between the projected area of fluorescent color and mean detection time (Figure 4):

$$y = 1.7526/x^{0.0902} \tag{1}$$

where

y = detection time (s) and
x = project fluorescent area (cm^2).

It is interesting to note that the helmet produced significantly faster detection times than many options with larger areas of fluorescent color. The reason for this is not known, but it could be surmised that the helmet shape is more easily associated with motorcyclists and hence reduces the detection time. There was strong evidence that detection times varied greatly depending on the nature of the site. Sites with large areas of unbroken color and low variations of light and shade resulted in relatively fast mean detection times. Slow mean detection times were found at busy sites where the amount of traffic produced a broken background pattern for visual search with numerous gaps and variations of color and shading, giving a patterned effect in which the motorcycle could be placed.

EFFECT OF HIGH-VISIBILITY AIDS ON DRIVER GAP ACCEPTANCE

It was felt that studying gap acceptance might prove fruitful as a field test of the effectiveness of conspicuity aids. It was hoped that the relative effectiveness of an aid could be related to changes in the observed distribution of gaps that motorists were prepared to accept in front of a motorcycle.

Study Design

It was decided to measure the gap-acceptance behavior of motorists toward a motorcycle in three conditions—control, dipped headlight, and fluorescent jacket. It was therefore necessary to introduce the experimental motorcycle into a traffic stream. It was apparent that to achieve an adequate rate of data collection the motorcycle would have to make repeated circuits past the junction in question. A rapid circuit was achieved by conducting the trials at a large roundabout, the Cock Pitt in Derby. The roundabout had a circumference of approximately 530 m (0.3 mile) with four access points; at the two junctions being studied, the path taken by the motorcycle was in the left-hand lane. A short pilot trial was conducted in which an experimental automobile preceded the motorcycle around the roundabout. The trials showed that gaps between 1.5 and 5.0 s would have to be used in the main trials to cover the range of accepted gaps, as suggested by Ashworth.

In the main series of trials, two videotape recorders were secured on a 3.6-m (12-ft) platform in the center of the roundabout. The trials were conducted by a team of six over a period of 4 d at the end of March 1977. The three options tested on the 250-cc motorcycle were

1. Control—The riders wore blue trousers; dark green jackets; black, open-face helmets; and black gloves.
2. Headlight—Conditions were the same as above except that the motorcycle headlight was switched on in the dipped condition (the lamp was 6 V and 24 W).
3. Fluorescent jacket—Conditions were the same as for item 1 but for the addition of a nylon fluorescent orange jacket.

The options were changed at half-hour intervals, and the order of presentation was varied between days to ensure even exposure to varying traffic conditions. In all, a total of 1854 passes were recorded on 10 half-hour tapes.

Video Analysis

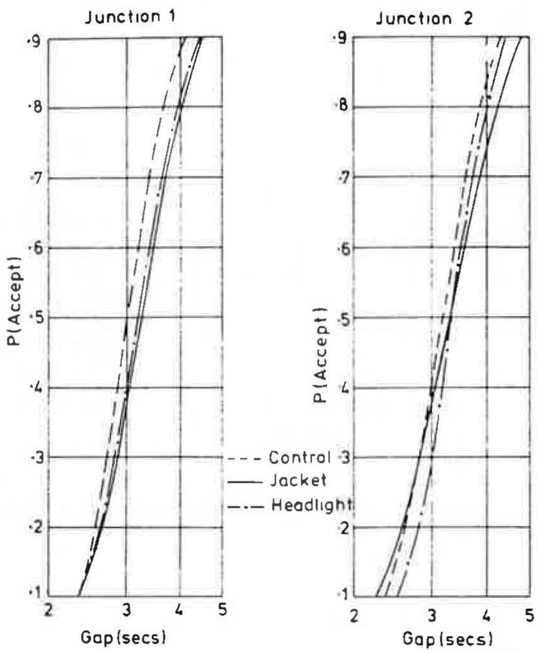
The video tapes were replayed on a Sanyo 1100 SL recorder and a Shibaden monitor at one-fifth real speed to permit tape analysis. Replaying at slow speed reduced errors in judgment when a vehicle passed a reference point and also reduced errors in reaction time in the analysis. The information taken directly from the tapes was the size of gap in seconds and, if the pass was valid, whether it was accepted or rejected.

Gaps were measured with a Colne electronic digital timer to the nearest hundredth of a second. The timer was started as the rear of the automobile passed a reference line and stopped as the front wheel of the motorcycle passed the line. A gap was included in the data only if it was a valid presentation, subject to the following criteria:

1. One or more automobiles or light vans had to be stationary at the junction as the lead experimental automobile passed by.
2. There was no interference from other traffic already on the island (i.e., passing the motorcycle and effectively shortening or filling the gap).

Sometimes a vehicle other than the experimental automobile preceded the motorcycle across the intersection. These gaps were measured and used in the analysis if the other vehicle kept to an acceptable line around the round-

Figure 5. Fitted curves for probability of gap acceptance at junctions 1 and 2.



about and the above criteria were satisfied. After all data had been taken from the tapes at one-fifth speed, they were all replayed at real time to check the accuracy of decisions concerning acceptance and rejection. At one-fifth speed it was sometimes difficult to judge whether some vehicles had come to a standstill at the junction before accepting a gap or if they had merely slowed down and then driven into the traffic stream. It was easier to make this classification when viewing at real speed. From the 1854 passes taped, a total of 352 acceptances and 922 rejections were recorded.

Data Analysis

The analysis of gap-acceptance data has been the subject of many papers (3, 4). The technique used here to analyze these data was the fitting of lognormal curves by probit analysis (5). Curves were also derived without the logarithmic transform, but the fit to the experimental data was poorer and the estimation errors on the median accepted gaps were much larger.

Results

The median accepted gaps and their 95 percent fiducial limits are given below:

Junction	Option	Median Accepted Gap (s)	95 Percent Fiducial Limits (s)
1	Jacket	3.25	2.96, 3.64
	Headlight	3.23	3.00, 3.50
	Control	3.07	2.87, 3.33
2	Jacket	3.31	2.79, 3.98
	Headlight	3.36	3.06, 3.76
	Control	3.21	2.87, 3.64

Figure 5 shows the cumulative distribution curves computed from the data. Clearly, the largest difference for the median accepted gap at either junction is only 0.18 s (between jacket and control conditions at junction 1), and there is considerable overlap of the limits on the medians. The median accepted gaps were compared for each junction. The largest difference was between the jacket and control conditions at junction 1, but this was not significant at the 0.05 level. Significant differences were not detected between any other medians. The slopes of the fitted lines corresponding to the standard deviations of the lognormal distributions did not differ significantly. The proportions of gaps of a particular size that were accepted in the noncontrol conditions were compared with the corresponding data for the control condition. No significant differences were obtained at either junction.

The analysis of the data from this series of field trials showed that the use of fluorescent clothing or a dipped headlight on the experimental motorcycle had no significant effect on the sizes of gaps accepted in front of it. The absence of any detectable change in the gap-acceptance behavior of motorists joining the traffic stream suggests that, if the motorcycle is perceived at the junction, the use of high-visibility aids has no effect on drivers' gap-acceptance behavior.

Although the presence of these high-visibility aids has not produced a detectable change in gap-acceptance behavior, it cannot be concluded that the use of such aids will have no benefit in the accident situation. The most important reason for the use of high-visibility aids is not to improve the drivers' perception of a motorcycle already detected but to ensure that the motorcycle is seen in the first place. On reflection, it seems unlikely that effects of this kind could be observed in an experiment studying gap-acceptance behavior.

The method in which the motorcycle followed the automobile around the traffic island was successful. It allowed rapid data collection in a natural traffic environment under controlled conditions. In addition, since it was unlikely that an observed gap was presented more than once to a vehicle waiting to enter the traffic stream, only one data point—an acceptance or a rejection—was recorded for each vehicle. Thus, the gap-acceptance functions obtained provide an essentially unbiased estimate of the population gap-acceptance response (6).

CONCLUSIONS

Several significant findings have come out of the work described in this paper:

1. Many pieces of high-visibility clothing have severe design problems and are strongly criticized by motorcyclists.
2. Most fluorescent materials show a strong tendency to lose both color and fluorescence in a relatively short time.
3. The time taken to detect a motorcyclist wearing a conspicuous color was shown to be inversely related to the projected area of color.
4. Neither the wearing of high-visibility clothing nor the daytime use of headlights affected motorist gap-acceptance behavior.

ACKNOWLEDGMENTS

We wish to acknowledge P. Watson and F. T. W. Lander of the Transport and Road Research Laboratory for their helpful advice and guidance throughout the project. This research was funded by the Transport and Road Research Laboratory. The views expressed in the paper are ours and do not necessarily represent those of the laboratory.

REFERENCES

1. P. Stroud and C. Kirkby. A Study of the Conspicuity of Motorcyclists in a Daylight Urban Environment. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, Internal Rept., 1976.
2. P. Stroud and C. Kirkby. Effect of Motorcyclists' High Visibility Aids on Driver Gap Acceptance. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, Internal Rept., 1977.
3. W. D. Ashton. Gap Acceptance Problems at a Traffic Intersection. Institute of Royal Statistical Society, Vol. 20, Pt. 20, 1971, pp. 130-138.
4. R. Ashworth. The Analysis and Interpretation of Gap Acceptance Data. Transportation Society, Vol. 4, 1970, pp. 270-280.
5. D. J. Finney. Probit Analysis. Cambridge Univ. Press, 1971.
6. R. Ashworth. A Note on the Selection of Gap Acceptance Criteria for Traffic Simulation Studies. Transportation Research, Vol. 2, 1968, pp. 171-175.

Discussion

Samuel P. Sturgis, Liberty Mutual Research Center, Hopkinton, Massachusetts

Ashford, Stroud, Kirkby, and Kirk have presented an extensive analysis of several important issues concerning the potential acceptability and effectiveness of methods for enhancing the daytime visibility of motorcyclists. They have found that a very small proportion of motorcyclists currently take the initiative of increasing their conspicuity by wearing high-visibility safety-related clothing. They have further found in a laboratory simulation that garments such as jackets and waistcoats of high-visibility colors can significantly decrease the time required for detection of a motorcyclist in an urban road environment.

It must be assumed that the observed reluctance of motorcyclists to wear high-visibility clothing stems from a belief that reported inconveniences associated with the clothing outweigh its possible usefulness in preventing accidents. A very important question that must be addressed from the standpoint of the motorcyclist then is, What is the role of conspicuity or lack of conspicuity in motorcycle accidents? This question is also of considerable importance to those who are concerned with evaluating the potential effectiveness of techniques for enhancing conspicuity.

Unfortunately, no direct answer to this question is available. However, an accident study conducted by Reiss, Berger, and Vallette (7) on a sample of motorcycle accidents that occurred in Maryland in 1973 does allow some inferences to be made. That study found that approximately 61 percent of motorcycle accidents involved collisions with other vehicles and that of these accidents 62 percent occurred at intersections. Reiss, Berger, and Vallette used a randomly selected sample of 200 such accidents, assigned culpability on the basis of accident descriptions by police, and found that the greatest single contributing cause was the failure on the part of the "other" driver to yield the right-of-way. This occurred in about 64 percent of the cases. Together, these percentages indicate that intersection accidents in which the other driver failed to yield accounted for approximately 24 percent of all (single and multivehicle) accidents studied. Reiss, Berger, and Vallette further found that, in the multivehicle intersection accidents, the motorcycle was most often proceeding straight ahead (86 percent of the cases) while the other vehicle was either turning left (49 percent), moving straight ahead (39 percent), or turning right (5 percent). The most common collision orientation involved the motorcycle striking the other vehicle at an angle (54 percent), and the next most common involved the other vehicle striking the motorcycle at an angle (21 percent).

Waller's 1972 analysis of the 630 multivehicle motorcycle accidents reported in North Carolina in 1968 (8) similarly concluded that culpability was attributable to the other driver in 62 percent of the cases. Waller further indicates that the predominant contributing circumstances in the multivehicle accidents studied were (a) the other vehicle turned in front of the motorcycle, (b) the other vehicle pulled out into the motorcycle, and (c) the other vehicle maneuvered without seeing the motorcycle. These categories accounted for 29, 20, and 10 percent of the accidents studied respectively.

Clearly, these studies indicate that drivers of other vehicles occasionally either do not perceive motorcycles, misperceive the location or speed of motorcycles, or intentionally fail to yield the right-of-way to motorcycles. It is not particularly surprising that these types of accidents occur at intersections since in many cases drivers entering an intersection must make very rapid decisions concerning the speed and location of vehicles approaching from several different directions. In addition, based on the relative number of motorcycle and other vehicle registrations in the United States, the probability of encountering a motorcycle rather than a larger vehicle on the road is relatively small. Thus, roadway encounters of automobile drivers with motorcycles are relatively rare events and as such are events that automobile drivers may not expect or specifically look for.

The implications of these findings for motorcyclists are quite clear: One should attempt to be as visible as possible and drive as defensively as possible, expecting occasionally not to receive the right-of-way when it is due.

These findings may also explain to some extent why greater differences were not found in the gap-acceptance study described by Ashford, Stroud, Kirkby, and Kirk in which drivers presented with a gap between an automobile and a motorcycle had only to contend with traffic approaching the intersections in question from one direction. As the authors point out, the primary purpose of high-visibility aids is to ensure that the motorcycle is seen in the first place. If the given detection task is too simple, one would probably not expect to find substantial differences in distributions of accepted or rejected gaps unless the sample sizes were extremely large. This may not be the case, however, in a more complex intersection situation where drivers are faced with traffic approaching from a number of directions.

Overall, the research presented is of considerable value to those concerned with the issue of motorcyle safety. It has shown that very few motorcyclists currently attempt to increase their conspicuity by wearing high-visibility clothing, that the styling and durability of many high-visibility garments is less than optimal, and that the use of high-visibility clothing can, at least in simulated conditions, significantly decrease the time required to detect a motorcyclist. Although the study did not find that the use of visibility-enhancing techniques had a measurable effect on the gap-acceptance behavior of drivers under the condition studied, it did show that the technique was procedurally workable and of potential value in future research on conspicuity.

REFERENCES

7. M. L. Reiss, W. G. Berger, and G. R. Vallette. Analysis of Motorcycle Accident Reports and Statistics. Motorcycle Safety Foundation, Final Rept., Feb. 1974.
8. P. F. Waller. An Analysis of Motorcycle Accidents With Recommendations for Licensing and Operation. Highway Safety Research Center, Univ. of North Carolina, Chapel Hill, 1972.

R. G. Mortimer, University of Illinois at Urbana-Champaign

In a large proportion of collisions between motorcycles and other motor vehicles, drivers of the other vehicles reported that they did not see the motorcyclist. This may or may not be the fact depending on the extent to which one is willing to accept reports of drivers who have been involved in such accidents. However, the geometric aspect posed by a motorcycle in many daytime driving situations and perhaps even more at night suggests that motorcycle and rider provide a target that is difficult to see.

Until the actual reason for these accidents is better understood, it is worthwhile to consider means of increasing the conspicuity of motorcycles and their riders. The study that is being discussed here was concerned with such an evaluation of the effectiveness of various aids to visibility in daytime conditions.

USER EVALUATION SURVEY

Apparently, 75.5 percent of the respondents to the opinion survey on safety-related clothing indicated that they were still wearing the clothing issued to them. This would suggest a generally high degree of satisfaction. That much of this clothing had an odd appearance was demonstrated by the fact that 25.3 percent of those participating in the study admitted to some degree of embarrassment in wearing it. This points out the need for good styling of clothing and integration of proper reflective materials into normal clothing worn by motorcyclists. Relatively few of the items that were evaluated in this study could be considered to be in the category of acceptably styled clothing that motorcyclists would willingly purchase.

Measurements of the degree to which the colors faded showed that the effectiveness of the clothing could not be assumed to be retained over very long periods of time, which indicates the need for improved materials.

This study should provide an impetus to the manufacturers of motorcyclists' clothing to make it better suited and more acceptable to motorcyclists and to provide improved visibility in daytime. Parenthetically, it would seem that an even greater effort needs to be made to ensure that clothing that is effective at night (9) should be more readily available.

LABORATORY SIMULATION OF EFFECTIVENESS OF HIGH-VISIBILITY AIDS

The tachistoscopic study of detection and recognition times of motorcyclists in a visual scene revealed that there appeared to be certain differences according to the types of clothing being worn. Primarily, the jacket and waistcoat produced significantly lower detection times than the control condition. In addition, the authors reported that the helmet produced shorter detection times than the control condition, but this finding was based on the dubious use of multiple t-tests. Although it was not stated by the authors, it is assumed that the sleeves, helmet, headlamp cover, and leg shields did not differ in their effect on detection time or differ from the control condition. However, I am also assuming that, since the mean detection times for these items of clothing were approximately the same as those for the control condition, they would as a group have had longer detection times than those for the motorcyclist wearing the waistcoat or jacket.

One might, therefore, argue with the use of these data in terms of a nonlinear equation that relates the area of clothing to detection time. Basically, Figure 4 could be indicated by two points that represent the central tendency of the detection times for the group consisting of the sleeves, headlamp cover, helmet, and leg shields and the central tendency of the other group consisting of the waistcoat and jacket.

This experiment was worthwhile and indicated that there were differences that were probably attributable to the various visibility aids that were evaluated by the 892 motorcyclists.

EFFECT OF HIGH-VISIBILITY AIDS ON GAP ACCEPTANCE BY DRIVERS

In the field test, three configurations were evaluated in daytime: the control condition, the dipped headlight, and the fluorescent jacket. The use of a roundabout (traffic circle) was ingenious in that it allowed very frequent gap-acceptance measures to be taken dependent only on the extent of the traffic flow on the roundabout. There were 352 gaps accepted and 922 rejected out of a total of 1854 passes; this indicates that in 69 percent of the passes traffic that involved some decision on the part of other drivers was present. The authors reported that there were no differences in the median accepted gap times that were attributable to the three motorcycle-and-rider display configurations.

It might be questioned whether median gap times are the most appropriate basis for comparison. Clearly, there is an increased likelihood of accidents if short gap times are accepted. Thus, an evaluation of, for example, the 10th percentile values of accepted gap times might be more relevant to an analysis of a potential hazard in the gap-acceptance judgments of other drivers. In Figure 5 it can be seen that the 10th percentile values of the three configurations at junction 1 are the same, whereas at junction 2 there is a spread in the gap times accepted for the three configurations that is greater than the spread between the medians. Thus, it appears that the headlight was somewhat more effective than the other two configurations in increasing the gap times accepted at the low end of the distribution. Whether such differences are significant has not been evaluated.

Although the authors conclude that it is quite likely that this type of experiment could not demonstrate any effect on the effectiveness of high-visibility aids whose function is to improve the detection of a vehicle, I do not feel this to be entirely the case. However, there might be another effect besides an effect on detection of using either the headlight or the fluorescent jacket. These aids may have increased the apparent image size of the motorcycle and its rider. If so, they could have had an effect not just on detection but also on the perception of distance and velocity. In that case, an effect on gap acceptance attributable to perceptual factors rather than increased likelihood of detection might have been noted.

It would also be interesting to evaluate whether the gaps accepted were discriminatory against the motorcyclist by using an automobile to make a comparison in the same situation of gap acceptance. This would help to answer questions such as whether or not the gaps that are accepted with respect to motorcycles are different from those accepted with respect to other vehicles for any number of reasons including perceptual as well as risk-taking factors.

In conclusion, it is felt that this research was most worthwhile, was carried out in a logical progression of studies concerned with various facets of the problem of motorcycle visibility, and used well-devised techniques to obtain the data. Obviously, more work needs to be done to improve detectability and provide other vehicle drivers with better information concerning the movements of motorcyclists.

REFERENCE

9. H. L. Woltman and R. L. Austin. Some Day and Night Visual Aspects of Motorcycle Safety. TRB, Transportation Research Record 502, 1974, pp. 1-8.

Paul L. Olson, Highway Safety Research Institute, University of Michigan

The authors are to be complimented for a high-quality, comprehensive piece of research. I would only like to comment on the problem to which the paper is addressed—motorcycle conspicuity.

The Highway Safety Research Institute (HSRI) is currently under contract to the National Highway Traffic Safety Administration (NHTSA) to investigate ways of improving the conspicuity of motorcycles. Thus, our program has aims quite similar to those described in the paper by Ashford, Stroud, Kirkby, and Kirk. Specifically, our charge was to conduct an analysis of motorcycle accidents, select promising conspicuity treatments, and carry out a field test program. Interestingly, the field test methodology we are using involves measures of gap acceptance.

Our analysis of the accident literature, based on about 10 000 accidents involving automobiles and motorcycles in the state of Texas in 1975, indicated that the precrash geometric relations were somewhat different than they were for accidents involving two automobiles. Notably, motorcycles tend to be involved in accidents when an automobile attempts to maneuver across their path. Specifically, the two situations that seem to be most significant in this respect are (a) what we have come to call a right cross or left turn and (b) a center-left turn. The former is a situation where the automobile is initially stopped on the right of the motorcyclist and is attempting to enter the roadway, either to cross completely or to perform a left merge maneuver. In the second situation, the automobile is initially facing toward the motorcycle and attempting to make a left turn across its path.

The overrepresentation of these two kinds of collisions in the motorcycle accident picture suggests that there is a problem with motorcycle conspicuity. We cannot be certain at this time what the exact problem is. It may be, for example, that the motorcycle is simply more difficult to see because it is considerably smaller than the bulk of the vehicles on the road. On the other hand, it may be that the motorcycle is seen but tends to be classified with pedestrians and bicycles whose mass it more nearly approximates. Whatever the reason, it appears that motorcycles would benefit from improved conspicuity and means of identification.

A variety of candidate conspicuity treatments were developed by using available materials. The various treatments were evaluated subjectively by a committee composed of NHTSA and HSRI personnel. Several of these were selected for initial field testing.

The first step in the field testing program was to determine whether the criterion selected was capable of discriminating among the various treatments. To do this, the first testing compared a control condition with several treatments that were very conspicuous; minimum regard was given to their appeal to the people who would have to use them.

As I mentioned earlier, a gap-acceptance methodology was employed in our study as well. It seemed clear to us as it apparently did to the authors of the paper being discussed that, if one can measure actual changes in the behavior of drivers maneuvering in front of a motorcycle, it is far more meaningful evidence of the effectiveness of a treatment than are the types of data provided in previous investigations. Obviously, if gap-acceptance measures show any changes, they imply that crashes arise from a fairly general response on the part of drivers and not, for example, from rare instances of poor judgment. Thus, negative results do not necessarily mean the treatments are ineffective.

I was impressed by the experimental method used in the gap-acceptance study described by Ashford, Stroud, Kirkby, and Kirk. It was a model of simplicity and good control. Unfortunately, if I understood it correctly, only one type of maneuver was possible for the automobiles. That maneuver would correspond (when corrected for the fact that Americans drive on the wrong side of the road) to what we call a right-right turn. This is not one of the maneuvers that our accident analysis suggests is particularly dangerous. For this reason we wanted to carry out our study in a way that allowed us to collect data on the two maneuvers described earlier (right cross or left turn and center-left turn). We did, however, collect data on the right-right turn maneuver as well.

Briefly, the data are collected in the following way. The motorcycle is driven along a busy thoroughfare in a city near Ann Arbor, Michigan. It is a very busy street with a great deal of cross traffic from shopping centers, restaurants, and so on. The speed limit is 72.5 km/h (45 mph). The motorcyclist is instructed to position the motorcycle behind a cluster of other traffic and to open a gap of about 100 m (a few hundred feet). As the experimenter rides along under this condition, he or she monitors traffic on the right and the left. If the motorcyclist sees a vehicle in position to make one of the three maneuvers of interest, he or she turns on the recording equipment with which the motorcycle is equipped and prepares to take data. The motorcycle is provided with equipment to measure distance traveled and an array of buttons to code various things. By pressing the appropriate buttons at the appropriate times, the experimenter can measure the size of the gap presented, report whether it was accepted or rejected, and what kind of maneuver was involved. These data are stored on magnetic tape and analyzed by computer.

We currently have data on five daytime conditions: (a) control motorcycle, (b) control automobile, (c) motorcycle equipped with a fluorescent fairing, (d) motorcyclist wearing a fluorescent jacket and helmet cover, and (e) motorcycle with low-beam headlight on. Not all of the conditions have as much data as we would like to see or will ultimately collect. The data we have at this time suggest that it may be possible to measure changes in driver behavior by the method described. It must be remembered that the study is in progress and conclusions at this time are tentative. We are encouraged by trends that show changes in the probability of acceptance of short gaps (less than 5 s) as a function of the treatment conditions investigated. However, these trends are only found in the right cross or left turn and center-left turn maneuvers. The maneuver that is most similar to that measured by Ashford, Strond, Kirkby, and Kirk seems to show no differences.

Again, I think this is an excellent paper. It is regrettable that the gap-acceptance methodology provided negative results, but it may be that an expansion of the technique will still prove meaningful.

Publication of this paper sponsored by Committee on Visibility.

Signalization of High-Speed, Isolated Intersections

Peter S. Parsonson, Georgia Institute of Technology

At signalized intersections where approach speeds are 56 km/h (35 mph) or higher, drivers face a "dilemma zone." If the yellow signal comes on while the driver is in this zone, a decision to stop may result in a rear-end collision or a sideswipe. The opposite decision, to go through the intersection, might produce a right-angle accident. For such an intersection, the traffic engineer needs to select a detector-controller configuration that will (a) detect an approaching vehicle before it enters the dilemma zone and either (b) extend the green signal to provide safe passage through the zone or else (c) end the green signal when the vehicle is still upstream of the dilemma zone and thereby provide adequate stopping distance. A major research project examined in detail a number of advanced detector-controller designs. The resulting design manual has systematically integrated into a single publication the available knowledge on the subject. This paper condenses the author's contribution to the design manual, elaborates on certain points incompletely treated by it, and proposes a new configuration. Current knowledge of dilemma-zone boundaries is reviewed, a classification of controllers and detectors and a taxonomy of detector-controller configurations are provided, and research data on the effectiveness of green-extension systems are summarized. The proposed new configuration uses a basic, actuated, nonlocking controller; 25-m (85-ft) long, delayed-call loop detector at the stopline; and two extended-call detectors upstream to give protection to the dilemma zone.

For over a decade, it has been known that at signalized intersections where approach speeds are 56 km/h (35 mph) or higher drivers face a "dilemma zone" or "zone of indecision." If the yellow signal comes on while the driver is in this zone, the decision whether to stop or go through may be difficult. A decision to stop abruptly may result in a rear-end collision. The opposite decision, to go through the intersection, might produce a right-angle accident. If the traffic-signal controller is vehicle-actuated rather than pretimed, the traffic engineer can attempt to design the installation to minimize the problem of the dilemma zone.

The goal of the traffic engineer in tackling this problem is to ensure, if possible, that no vehicle is in the dilemma zone on the display of the yellow interval. The key to the solution is the selection of a cost-effective detector-controller configuration that will (a) detect an approaching vehicle before it enters the dilemma zone and either (b) provide safe passage through the zone or (c) provide adequate stopping distance. Thus, the solution focuses on the placement of vehicle detectors and the coordination of that placement with the timing functions of the controller.

It bears emphasizing that the dilemma zone can be protected only if the green signal is terminated by "gap-out." If the green is extended by heavy traffic (or an overlong unit extension) to the maximum interval, there can be no protection. A vehicle may well be caught in the dilemma zone.

A major research project examined in detail a number of advanced detector-controller designs for use at high-speed, isolated intersections. The resulting design manual (1) systematically integrated into a single publication the available knowledge on this subject. This paper condenses my contribution to the design manual and elaborates on certain points incompletely treated by it. A new configuration is proposed.

The dilemma caused by indecision on the display of the yellow interval is the subject of this paper but is only one of three separate difficulties associated with the termination of the green interval. A second and different dilemma faces the motorist if the length of the yellow interval (plus any all-red clearance interval) is not enough to permit him or her either to clear the intersection or to stop safely (2). A third type of dilemma is the "short green" problem in high-speed signalization (3). A green interval of only 2 to 4 s in length may so conflict with a driver's expectations that he or she may panic and not react to the yellow change interval although there is ample opportunity to stop.

BOUNDARIES OF THE DILEMMA ZONE

Once it has been determined from analysis of accidents or conflicts that the problem of a dilemma zone exists on an approach, despite a rational timing of the yellow-plus-all-red clearance period, an advanced detector-controller configuration is warranted. The first step in the selection of this configuration is the identification of the extent, or boundaries, of the dilemma zone. This can be obtained from the literature and adjusted for gradients (4).

In 1974, Parsonson and others (5) examined research on the probability of stopping from various speeds (6, 7, 8). They characterized the dilemma zone as that approach area within which the probability of deciding to stop on the display of yellow is within the range of 10 to 90 percent. That is, the upstream boundary of the dilemma zone was located where 90 percent of drivers would decide to stop if the yellow began just as they reached that boundary. At the downstream boundary, closer to the stopline, only 10 percent would decide to stop. These findings are summarized below (1 km/h = 0.62 mph and 1 m = 3.3 ft):

Approach Speed (km/h)	Distance From Intersection for Two Probabilities of Stopping (m)	
	10 Percent	90 Percent
48	27	52
64	33	75
72	50	90
80	66	105
97	78	135

These data agree well with data for 48 and 80 km/h (30 and 50 mph) published by Olson and Rothery in 1972 (9).

Zegeer of the Bureau of Highways, Kentucky Department of Transportation (DOT), recently conducted a thorough study of dilemma-zone boundaries for nine straight and level approaches (4). Responses of about 2100 drivers to the yellow interval were recorded. Figure 1 and the following table summarize Zegeer's findings:

Approach Speed (km/h)	Distance From Intersection for Two Probabilities of Stopping (m)	
	10 Percent	90 Percent
56	31	77
64	37	86
72	46	99
80	52	107
89	71	117

It can be shown that at speeds of 72 to 80 km/h (45 to 50 mph) Zegeer's dilemma zones are 28 to 38 percent longer than those reported by Parsonson and others (5). At lower and higher speeds, the differences are minor. The Zegeer data are extensive and were collected under closely controlled conditions. Most traffic engineers will probably use his findings in Figure 1 and the table above rather than data given in the earlier table.

The Zegeer data show that the upstream boundary of the dilemma zone, at which 90 percent of motorists will decide to stop, is 4.5 to 5 s of passage time from the intersection. The other boundary, for a 10 percent chance of stopping, is 2 to 2.5 s from the intersection. There is a dilemma zone that is typically 2.5 to 3 s in length.

Any solution to the problem of the dilemma zone begins with the detection of an approaching vehicle before it enters the dilemma zone. Therefore, it is axiomatic that there should be a detector approximately 5 s of travel time before the stopline, just upstream of the dilemma zone. In this connection, it is useful to show the dilemma-zone "cloud" (shaded area) on a table of approach speed versus passage time from detector to stop line (Figure 2) (4). The figure shows that 5 s of detector setback is adequate except for speeds of 97 km/h (60 mph) or more. Cell values are distances in meters from the detector to the stopline at that approach speed.

In the years before there was wide circulation of research data on dilemma-zone boundaries, it was common for traffic engineers to derive the boundaries from kinematic analyses of stopping and clearing. The upstream end of the dilemma zone is associated with stopping; therefore, a calculation of safe stopping distance from a certain approach speed should give a satisfactory estimate of the correct location for a detector just upstream of the dilemma zone. The minimum stopping sight distances for wet roads of the American Association of State Highway Officials (AASHO) (9) are shown in Figure 2 (dashed line) to be reasonably close to the upstream boundary of the dilemma zone. Figure 1 indicates a high probability of stopping (96 percent) for 80 km/h (50 mph) and the AASHO safe stopping distance of 111 m (369 ft). A detector placed at this location would lay the groundwork for excellent protection against dilemma.

Some investigators have not used the AASHO safe stopping distances but have instead assumed a 1-s reaction time and an emergency stop on a dry road. Bierele (10), Grimm (11), and, in personal correspondence, Holloman, assistant traffic engineer for the city of Winston-Salem, North Carolina, have reported stopping distances at 80 km/h of approximately 76 m (250 ft) on this basis of calculation. Figure 1 indicates a probability of stopping of only 47 percent for 76 m and 80 km/h. A detector placed at this location would leave over half of the dilemma zone without detection.

DETECTOR-CONTROLLER CONFIGURATIONS

The purposes of this section are (a) to establish a uniform terminology, (b) to organize a taxonomy of advanced detector-controller strategies, and (c) to propose a simple, qualitative flow chart to assist the traffic engineer to sort out the alternative strategies.

Terminology

Sackman and others (1) explain the meaning of many specialized terms, such as stretch detector, locking detection memory, and modified density controller. The distinction between several specialized types of detectors is important to this paper.

Here, the term extended-call detector is used to describe a unit that has a carryover output: When the vehicle leaves the detection area, the extended-call detector "stretches" or prolongs the call for an adjustable period of seconds. An extended-call detector connected to a small loop or single magnetometer probe can essentially mimic the output of a normal detector connected to a very long loop or a series of probes.

By contrast, a delayed-call detector does not issue an output until the detection zone has been occupied for an adjustable number of seconds. Delayed-call detectors are finding extensive use in detecting congestion (5) and in screening out false calls for the green signal (3).

In this paper, a green-extension system is a unit offered by at least two manufacturers to provide protection for the dilemma zone at a semiactuated intersection (5). The unit includes two or more extended-call detectors and also display-monitoring circuits that aid in the control of the end of the green.

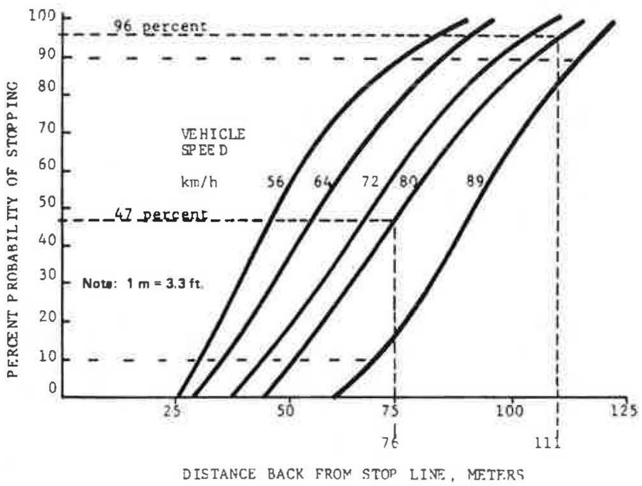
Table 1 gives a taxonomy of detector-controller configurations. It systematizes the advanced designs currently in use, or advocated for use, in the United States. Each design is "advanced" in that it uses multiple-point detection or advanced actuated controller or both. Details of the applications of these designs can be found elsewhere (1). Table 1 demonstrates how different agencies and engineers have combined various components of detector-controller hardware in their attempts to achieve safety at high-speed intersections. The table covers all types of controllers and is careful to distinguish between basic and "density" models and locking and nonlocking detection-memory modes. Similarly, detection is clearly specified as to type.

Figure 3 is a flow chart intended to assist the traffic engineer to make a preliminary selection from the detector-controller configurations given in Table 1. There are several key questions on the flow chart that the traffic engineer needs to be able to answer for purposes of specific application. The first is, Is it important for efficiency that the equipment also be capable of changing the green on detection of a gap no greater than 2 to 4 s? If the answer is no and the traffic engineer is willing to accept a gap of 5 s, then the flow chart leads to relatively simple designs that use basic actuated controllers and detection systems that are comparatively inexpensive. Many traffic engineering agencies in the South and the Southeast, for example, find that these designs are adequate for their needs. If the answer to the question is yes (as, for example, in many jurisdictions in the West Coast states), then the flow chart leads to relatively complex designs that use density controllers or extensive detectorization or both. The degree of complexity and expense in these categories is primarily a function of whether traffic conditions are so variable during a 24-h period that the equipment must be able to measure speed and, in response, change the control logic in real time.

For those engineers who answer no to the question above, the next key question is, Are false calls for the green (as in right turn on red) numerous enough that it is important that the equipment have at least some capability to screen out false calls? Practically every jurisdiction in the United States permits right turn on red in some form. The capability to screen out false calls is so vital to the efficiency of any actuated intersection, urban or rural, that it would seem that most traffic engineers would answer yes to this question. The flow chart will then indicate a basic, fully actuated, nonlocking controller with a long presence loop at the stopline and an extended-call detector to protect the dilemma zone. If the screening out of false calls is of particular importance—a major goal—then a refined design that includes a delayed output from the stopline loop is suggested.

The flow chart does not venture into the area of maintenance of controllers and detectors. It is left to the traffic engineer to factor in such important considerations as the capability of maintenance staff and the difficulty of keeping detection loops in service.

Figure 1. Dilemma-zone curves for Kentucky drivers.



KINEMATIC ANALYSES OF SELECTED CONFIGURATIONS

It is useful to analyze the various designs given in Table 1 by posing certain questions, most of which require kinematic analyses:

1. Does the design detect a vehicle approaching at the design speed before it reaches the dilemma zone?
2. What is the allowable gap imposed by this design? The allowable gap is the maximum time interval between actuations that will cause the green to hold. A short al-

Figure 2. Dilemma-zone boundaries.

APPROACH SPEED, KILOMETERS PER HOUR	PASSAGE TIME IN SECONDS FROM DETECTOR TO STOP LINE: 1	2	3	4	5	6
48	15	27	40	54	67	81
56	16	31	47	62	78	93
64	18	36	54	72	90	108
72	20	40	60	81	101	121
80	22	45	67	89	111	134
89	25	49	74	99	124	148
97	27	54	81	107	134	161
105	29	58	87	116	145	174

Note: 1 km/h = 0.62 mph and 1 m = 3.3 ft.

lowable gap will cause the green to terminate in a "snappy," traffic-responsive manner. A long allowable gap will often prolong a green until it is terminated by the maximum interval setting of the controller. This is highly undesirable because no dilemma-zone protection is provided on "max-out" and a vehicle may well be caught in the dilemma zone.

3. On termination of the green by gap-out, will the vehicles approaching at the design speed be clear of the dilemma zone?

4. On termination of the green by gap-out, will vehicles traveling slower than the design speed be clear of the dilemma zone?

5. Can a queue waiting at the stopline get into motion without a premature gap-out?

6. Can the design screen out false calls for the green (as, for example, in right turn on red)?

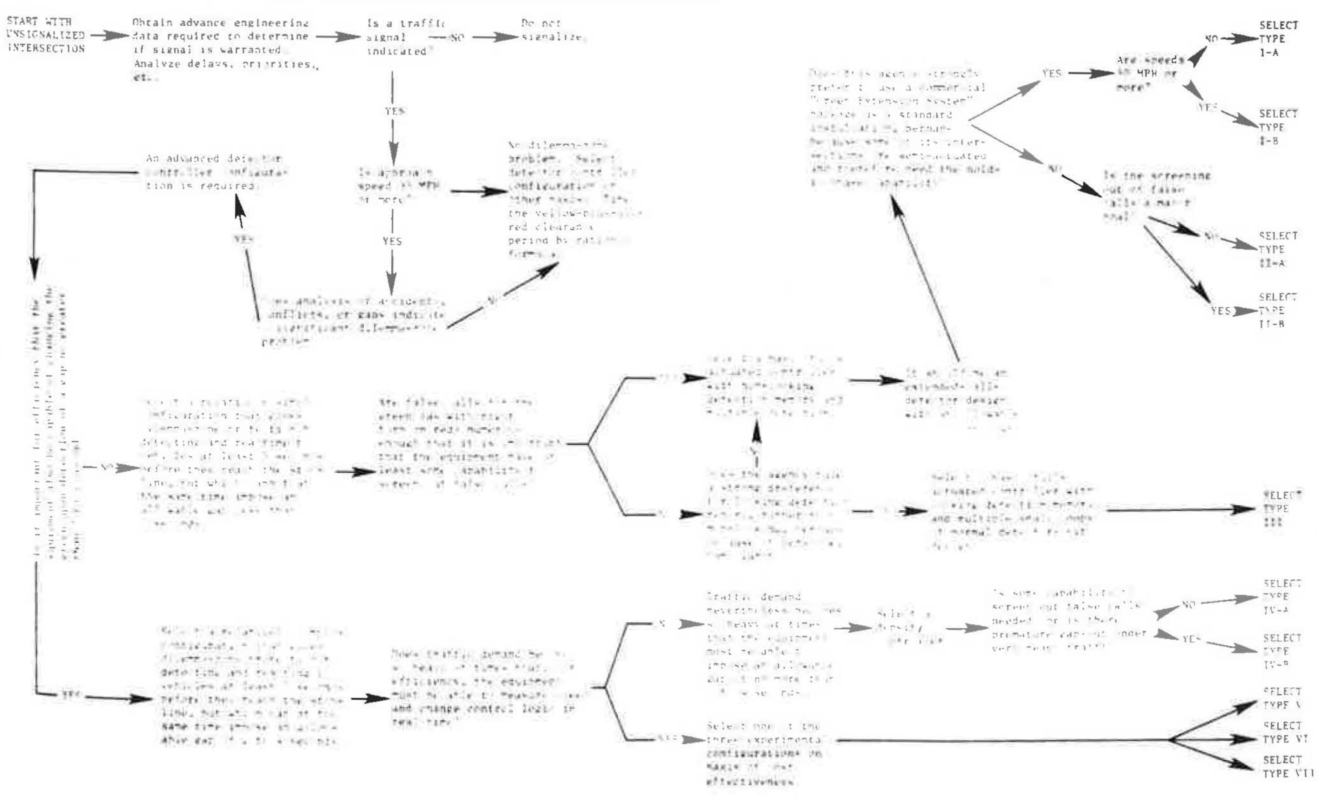
7. During the green interval, can a queue of left-turning vehicles hold the green as they wait to filter through gaps in oncoming traffic? This is important on

Table 1. Taxonomy of detector-controller configurations.

Type	System	Design	Use
1	Green-extension systems for semi-actuated controllers	Two-loop Three-loop	Composed of extended-call detectors and auxiliary logic: controller normally semiactuated with either locking or nonlocking memory; green-extension systems can also be used with basic, fully actuated, nonlocking controllers, in which case auxiliary logic is not needed
2	Extended-call detection systems for basic controllers	21-m loop at stopline (normal detector output) supplemented by extended-call detector 5 s before the stopline 21-m loop at stopline (delayed-call output) supplemented by extended-call detector 5 s before the stopline and a third detector (of normal or extended output) between them	Used with basic, fully actuated, nonlocking controllers
3	Multiple-point detection system for basic controllers		Composed of multiple small-area detectors positioned to take into account vehicles traveling at and under design speed; used with basic, fully actuated, locking controllers
4	Systems for density controllers	Conventional design using one small-area detector located 5 s before the stopline Extended-call detection systems	
5	Shifting-presence zone detection systems		Composed of many hypothetical speed-detection sensors, each sensitive to a narrow speed range and positioned to maintain green for a wide range of approach speeds; intended for use with density controllers
6	Area detection system with volume adjustment		Composed of 18-m presence loop at the stopline, supplemented by multiple-point detection for a distance of 244 m; volume-level indicator disconnects upstream detectors when heavy volumes indicate lower speeds; uses a basic controller
7	Computer controller		Computer measures speed of each vehicle and continuously adjusts the vehicle extension interval to provide dilemma-zone protection at all speeds

Note 1 m = 3.3 ft.

Figure 3. Flow chart for preliminary selection of detector-controller configurations.



two-lane roads, where an occasional left-turning vehicle can cause a queue to form. When a gap in oncoming traffic appears, a gap-out may occur before the queue can get into motion over its detector.

These criteria are applied here to one of the configurations in Table 1 for an example design speed of 72 km/h (45 mph). The table given previously for Zegeer's data (4) gives the dilemma-zone boundaries for this speed as 46 and 99 m (152 and 325 ft) from the stopline, which corresponds to 2.3 and 4.9 s of passage time respectively.

Conventional High-Speed Design

The conventional or most straightforward design for a 72-km/h (45-mph) signalized approach would use a single small-area detector 99 m (325 ft) before the intersection. The controller would be an advanced actuated model (density or volume-density) with the following key timing settings: passage time, 5 s; minimum gap, 3 s; and last-automobile-passage feature, disabled if present. Figure 4 shows the detector location and the dilemma zone.

The answers to the seven questions posed previously are as follows:

1. The design does detect a design-speed vehicle before it reaches the dilemma zone since the upstream detector was located in accordance with Zegeer's data on dilemma-zone boundaries, given previously.

2. The allowable gap imposed by this design reduces, usually on the basis of time waiting on the red, to the setting of the minimum gap (here 3 s). The shortest setting that would pass a 72-km/h (45-mph) vehicle through a 53-m (173-ft) dilemma zone is 2.6 s. The 2.6 s is therefore the minimum desirable allowable gap; a shorter value would give snappier operation but could leave a vehicle in the dilemma zone.

3. On gap-out, the vehicles will be clear of the dilemma zone, and the last automobile to have crossed the detector will be 3 s downstream from it and 2 s from the stopline. The driver will have little difficulty in deciding to go through. The vehicle behind the last automobile (termed here the trailing automobile) will be upstream of the detector on gap-out and will easily decide to come to a stop. A typical positioning of vehicles on gap-out is shown in Figure 5.

If the controller incorporated the feature for last-automobile passage, the signal indication would not change until the last automobile had completed its passage time of 5 s after crossing the detector. Figure 6 shows that in this case the trailing automobile may well be caught in the dilemma zone. The figure explains why California, for example, does not use the last-automobile-passage feature.

4. On termination of the green by gap-out, vehicles traveling slower than the design speed may be caught in the dilemma zone if the allowable gap is set low. For example, a minimum allowable gap of 4.9 - 2.3 = 2.6 s barely permits a vehicle at the design speed of 72 km/h (45 mph) to clear its dilemma zone before gap-out. A slower vehicle would be caught. Figure 7 shows that an allowable gap of 4.3 s would be required to pass a straggler at 56 km/h (35 mph) through its own dilemma zone. It can be seen that there is a trade-off between snappy operation and protection for the slower vehicles in the traffic stream. One can be obtained only at the expense of the other. What is needed is a detector-controller configuration that can measure the speed of the last automobile and tailor an appropriate extension of the green. Computer controllers can do just that (see type 7 in Table 1) and represent one alternative to the conventional design.

5. A queue waiting at the stopline is supposed to be able to get into motion without premature gap-out. A density controller has a "variable initial interval," which is intended to produce a minimum green sufficient to permit motion over the detector in time to extend the green. Thus, the design taken at face value will meet this test. However, in practice it has been observed that premature gap-out can occur when traffic is very heavy. Dense traffic can defeat the purpose of the timing adjustment that controls the number of actuations (on the red) that will cause maximum initial to time. When traffic is dense, traffic stopping at the intersection may arrive at the detector during the green interval and therefore contribute nothing to the timing of the next initial interval. Years ago, the only remedy was to set a value of minimum initial high enough to ensure motion over the detector. Such a high setting resulted in slug-

Figure 4. Conventional detector-controller design for 72-km/h (45-mph) approach speed.

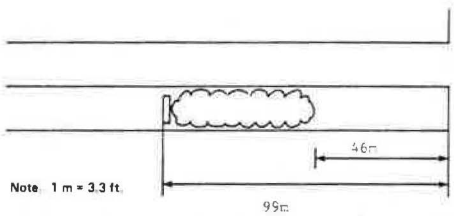


Figure 5. Typical positioning of last automobile and trailing automobile on gap-out without last-automobile-passage feature (conventional high-speed configuration).

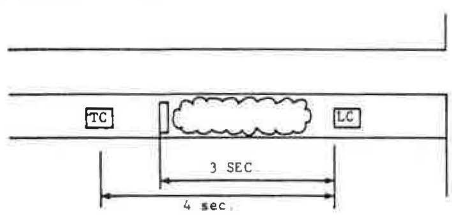


Figure 6. Conventional high-speed detector-controller design with last-automobile-passage feature.

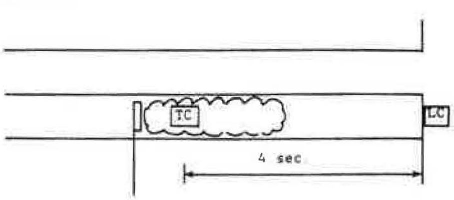
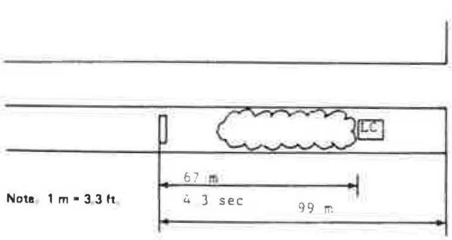


Figure 7. Gap-out of 56-km/h (35-mph) vehicle in the conventional high-speed design for 72 km/h (45 mph).



gish operation during periods of light traffic and a loss of reputation for the sophisticated capability of the density controller. The type 4 extended-call system offers a solution to this problem and is discussed elsewhere (1).

6. The design has no ability to screen out false calls for the green because the controller's detection memory is of the locking type. Once a vehicle crosses a detector, its call will be locked in until satisfied by a display of the green to that approach even if the vehicle has turned into a gas station or turned right on red. Many of the alternative designs in Table 1 provide a degree of screening.

7. A queue of vehicles waiting to turn left cannot hold in a call for the green. Many of the alternative designs in Table 1 overcome this problem by using a stopline loop.

Green-Extension Systems for Semiactuated Controllers

A green-extension system (GES) is a commercially available equipment package consisting of two or more extended-call detectors, one or more auxiliary timers that can disconnect or "force off" the extended-call detectors, and auxiliary electronics that can monitor the signal display, arm or make operational the extended-call detectors, and control the yielding of the green to the side street (by activation of hold-in-phase circuits). The auxiliary timers and electronics are needed only if the controller is semiactuated. If it is fully actuated, then the extended-call detectors do not require any auxiliary logic and the designs are as given for type 2 (Table 1). The semiactuated controller can use either locking or nonlocking memory for the side street depending on whether detection is for a small or large area.

The type 1 two-loop system uses two extended-call detectors and is considered satisfactory for approach speeds up to 72 km/h (45 mph). The three-loop systems are recommended where approach speeds are in excess of 72 km/h or where speeds are lower but traffic densities are quite high. The allowable gap of a GES is typically 4.5 to 5 s.

Inasmuch as Parsonson and others (5) describe GESs in detail, and since semiactuated control is steadily losing favor for use at isolated intersections, no further consideration of such systems is required here.

EFFECTIVENESS OF GREEN-EXTENSION SYSTEMS

There is a substantial amount of before-and-after data on the effectiveness of GESs in Kentucky. Zegeer has prepared an outstanding report on the effectiveness of 5 of 16 GES installations of the Kentucky DOT (4). Extensive accident data for 3 of these sites were combined to give a total of 8.5 years of before data and 3.7 years of after data. Zegeer found a total of 70 accidents before GES and 14 accidents after or 8.2 and 3.8 accidents/year respectively. This was a reduction of about 4.4 accidents/year or 54 percent. Zegeer reported that rear-end accidents were reduced by 75 percent and right-angle accidents by 31 percent. Summaries of property-damage-only, injury, and fatal accidents showed that the number of each type of accident was reduced by approximately 50 percent.

Two new GES sites, at the Kentucky towns of Ashland and Stanford, were selected for before-and-after studies of conflicts, speeds, and delays. Average speed at both intersections is approximately 66 km/h (41 mph), and each uses two-phase, semiactuated control. Figure 8 (4) shows as an example the installation at the Ashland intersection. The five detectors shown on US-23 are GES detectors; they do not actuate the controller. The 4 percent downgrade on the northbound approach determined the need for a third GES detector 125 m (410 ft) from the stopline. The comprehensive evaluation of the two intersections produced a number of significant conclusions, including the following:

1. The six types of yellow-phase conflicts observed were reduced by an average of 62.1 percent.

2. No significant change was found in the number of automobiles stopped or in the total delay of vehicles on side streets after installation of the GES.

3. The initial cost to install a GES to an existing signal is $2750, and maintenance costs for a 10-year period are $500/year. The cost of an average accident to the highway user in Kentucky is $7112. Therefore, if a GES installation were to eliminate only one mainline, rear-end accident per year, the benefit/cost ratio would be 6 and the total net benefit to motorists would be close to $30 000 over a 10-year period.

PROPOSAL FOR A NEW CONFIGURATION

There appears to be an unmet need for a high-speed design that includes loop-occupancy features; a basic, actuated, nonlocking controller; and extended-call detectors and that provides both a short allowable gap and protection over a wide range of speeds. A new configuration of the type 2 delayed-call variety is proposed in this section and is shown to have an allowable gap of 3.3 to 4.0 s and a range of speeds from 56 to 80 km/h (35 to 50 mph).

Figure 9 shows the details of the design. The 26-m (85-ft) long stopline loop is a delayed-call design [with a quadrupole (3) configuration to improve detection of small vehicles]. So great a length is intended to hold the call of discharging vehicles until a 2-s gap in 56-km/h (35-mph) traffic occurs. In this way, the green will be held by start-up traffic until motion over the extended-call detectors is ensured. Premature gap-out is thus avoided.

The following analysis presumes that (a) both extended-call detectors are the type that time the extension from the exit of the vehicle and not its entrance into the detection area, (b) 1.8-m (6-ft) long loops are used, and (c) the vehicle is 4.5 m (15 ft) long.

The upstream detector is set to extend the call by 1.4 s. This is sufficient to carry vehicles at 64 to 80 km/h (40 to 50 mph) to the second extended-call detector (Figures 9 and 10). Slower vehicles at 56 km/h (35 mph) will not reach that detector, thereby losing their green before reaching their own dilemma zone (Figure 11). The second detector is set to extend the call by 1.9 s, to carry vehicles at 64 to 80 km/h through their dilemma zones. A kinematic analysis follows.

1. The design does detect an 80-km/h vehicle before it reaches the dilemma zone since the upstream detector was located in accordance with Zegeer's data.

2. The allowable gap is nominally the sum of the settings of the two extended-call detectors or 3.3 s. More precisely, the allowable gap should be calculated by taking into account the lengths of the loops and vehicles. On this basis, the time headway from front bumper to front bumper that will just hold the green is 3.7 s for 80-km/h traffic and 4.0 s for a 64-km/h stream. The fact that the stopline loops disconnect once discharging traffic is at speed is of great value in ensuring a reasonably short allowable gap.

3. On gap-out, vehicles traveling at the design speed—80 km/h—will be clear of their dilemma zone (Figure 9).

Figure 8. Intersection of US-23 and Hoods Creek Pike in Ashland, Kentucky.

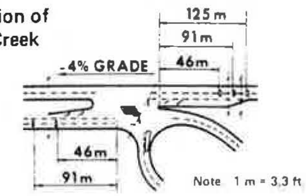


Figure 9. Proposed type 2 delayed-call design for 80-km/h (50-mph) approach speed.

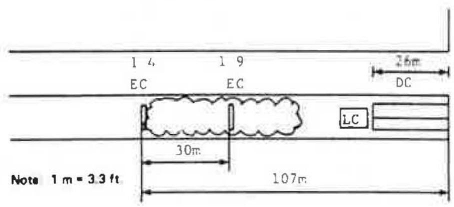


Figure 10. Gap-out of 64-km/h (40-mph) vehicle in the proposed design.

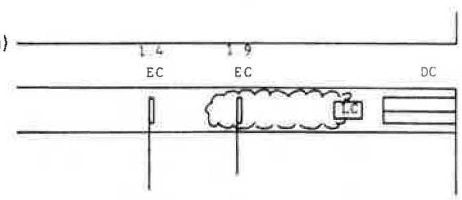


Figure 11. Gap-out of 56-km/h (35-mph) vehicle in the proposed design.

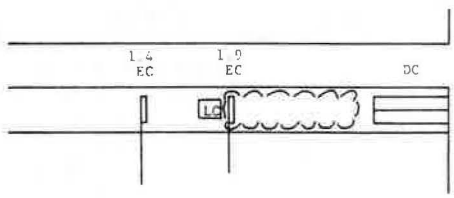
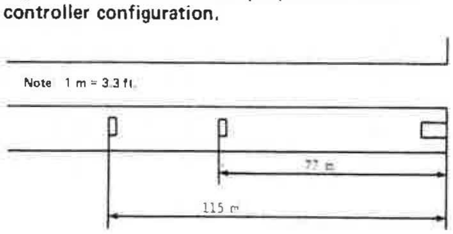


Figure 12. Modification of proposed detector-controller configuration.



4. Slower vehicles, at 56 to 72 km/h (35 to 45 mph), will also be clear of the dilemma zone on gap-out. See Figure 10 for 64 km/h and Figure 12 for 56 km/h.

5. The 26-m (85-ft) long loop at the stopline will allow a queue waiting at the stopline to get into motion without premature gap-out.

6. The delayed-call design of the stopline loop improves the ability of the design to screen out false calls for the green. When the green is at rest on the cross street, however, a false call at either of the extended-call detectors will bring the green unnecessarily.

7. The long loop at the stopline permits a queue of left-turning vehicles to hold the green as they wait to filter through gaps in oncoming traffic.

The next step in the development of this proposed new configuration will be a trial installation in the Atlanta area.

ACKNOWLEDGMENTS

Much of this paper is adapted from my contribution to a Federal Highway Administration research contract. I served as a consultant to the contractor, Urban Interface Group of Laguna Beach, California.

I am grateful to Charles V. Zegeer of the Kentucky DOT for permission to report his research results to the Transportation Research Board. The following individuals offered detailed, constructive, thoughtful comments on the first draft of this paper: James Gawlas of the Gwinnett County, Georgia, Traffic Engineering Division; Donald E. Lepic of the Adams County, Colorado, Community Development Division; Lionel M. Rodgers of Storch Engineers, Wethersfield, Connecticut; and Richard C. Wisler of Sarasota Engineering Company, Sarasota, Florida. Their suggestions are gratefully acknowledged. The manuscript was edited and typed by John C. Whatley of the School of Civil Engineering, Georgia Institute of Technology.

The contents of this paper do not necessarily reflect the official views or policies of the U.S. Department of Transportation.

REFERENCES

1. H. Sackman and others. Vehicle Detector Placement for High-Speed, Isolated, Traffic-Actuated Intersection Control: Volume 2—Manual of Theory and Practice. Federal Highway Administration, Rept. FHWA-RD-77-32, 165 pp., May 1977.
2. A. D. May, Jr. Clearance Interval at Traffic Signals. HRB, Highway Research Record 221, 1968, pp. 41-71.
3. P. S. Parsonson and others. Large-Area Detection at Intersection Approaches. Traffic Engineering, June 1976, pp. 28-37.
4. C. V. Zegeer. Effectiveness of Green-Extension Systems at High-Speed Intersections. Bureau of Highways, Division of Research, Kentucky Department of Transportation, Lexington, Research Rept. 472, May 1977.
5. P. S. Parsonson and others. Small-Area Detection at Intersection Approaches. Traffic Engineering, Feb. 1974, pp. 8-17.
6. R. Herman and others. Problem of the Amber Signal Light. Traffic Engineering and Control, Vol. 5, Sept. 1963, pp. 298-304.
7. F. Webster. Progress of Work on a New Type of Controller for Traffic Signals on High-Speed Roads. British Road Research Laboratory, Research Note RZ/3634 FVW, 1959.
8. P. L. Olson and R. W. Rothery. Deceleration Levels and Clearance Times Associated With the Amber Phase of Traffic Signals. Traffic Engineering, April 1972, pp. 16-19.
9. A Policy on Geometric Design of Rural Highways. AASHO, Washington, DC, 1965, p. 136.
10. H. Bierele. A Method of Detector Placement. IMSA Signal Magazine, May-June and July-Aug. 1974.
11. R. P. Grimm. Traffic Signal Operation Design: Efficiency and Safety. Newsletter of Western ITE, Oct.-Nov. 1974, pp. 2-4; ITE, Technical Notes, Vol. 1, No. 1, Summer 1975, pp. 2-4.

Discussion

Jon D. Clark, Kentucky Department of Transportation

Parsonson has provided a very valuable tool to the engineer whose objective is to design a signal system that will provide dilemma-zone protection for high-speed vehicles approaching an isolated signalized intersection. Many jurisdictions, recognizing the dilemma-zone problem, have developed and implemented unique system designs. Quite often these systems were designed on a case-by-case basis, and very little attention was given to standardization.

This paper and the research project design manual referred to by Parsonson (1) have analyzed and classified the various state-of-the-art and classic systems in use today and have provided a taxonomy of advanced detector-controller strategies for the practicing engineer. This taxonomy does provide a means of standardizing design as well as providing a common basis for future discussion for the practicing engineer.

The state of Kentucky uses a standardized design for the 32 dilemma-zone signal systems currently in operation. The system design used would most closely fit Parsonson's type 1 (two-loop and three-loop) green-extension classification even though basic, fully actuated, nonlocking controllers are used. A more complete description of the system design used in Kentucky can be found in the appendixes of the research project design manual (1) and a report by Zegeer (13).

It should be noted that practically all dilemma-zone protection signals in Kentucky are located on major arterials and were installed under the interruption of continuous traffic warrant. Capacity is seldom a major problem even though every effort is made to maintain a high level of operational efficiency. All 32 intersections currently provide dilemma-zone protection to the mainline phase only.

Originally, stop-bar loops were used for all approaches that had dilemma-zone protection. The initial interval was low and the efficiency high; however, from a safety standpoint, this type of operation proved to be less than desirable. This design, during off-peak periods, created unreasonably short mainline green periods, which in turn created an intolerable stopping problem, particularly for commercial vehicles. Time-lapse photography showed that commercial vehicles accelerated just before their arrival at the normal dilemma zone, particularly after they observed that the green phase commenced as they were approaching the intersection [303 m (1000 ft) or more]. Eliminating the presence loops and placing the phases on minimum recall with a 12- to 18-s initial interval, in addition to the advance loop extension time, seemed to satisfy driver expectations and lessen the problem. A second consideration for this minimum recall type of operation was the desire that the signal dwell in the dilemma-zone green phase during periods of rest or vehicle inactivity. This is very important when the dilemma-zone approach is on a significant grade and snow and ice are not uncommon.

Truck (commercial) traffic creates a very severe problem at several locations. This is particularly true at locations that have a significant downgrade approach where truck speed is excessive and sight distance is very good [0.8 km (0.5 mile) or more]. Automobiles share this problem to a lesser degree. The problem of the short green phase mentioned earlier becomes very significant under these conditions. Truck drivers, with their vantage point and experience, probably have a perception-reaction time 50 percent less than that of drivers of passenger vehicles; however, the actual stopping distance for trucks is much greater. For example, a vehicle with a 3-S2 commercial classification would require a 65 percent greater stopping distance than would an average vehicle (13). Excessive weight and poor brake performance can increase the stopping distance even more. Ideally, a truck detection system would be used and trucks would be treated separately. Unfortunately, this type of equipment is not currently available.

The technique used in Kentucky for an intersection with the aforementioned problem is to use the highest observed vehicle speed as the design speed in determining the location of the back loops. This technique has produced very satisfying results. An example is the Ashland (US-23) intersection (4, 12). The northbound steep-grade approach with a two-loop configuration (85th percentile design speed) experienced a much higher incidence of traffic conflicts than did the southbound approach, which was a 0 percent grade that used the same detection scheme. In an attempt to reduce the problem, an additional loop was installed on the northbound approach by using the 99th percentile speed as the design speed. This additional loop in the northbound direction reduced the number of conflicts to the same approximate level as that for the southbound approach. The results of a conflict study conducted at this intersection by Zegeer (4) are given below:

Condition	Conflict Rate per 1000 Opportunities (%) North-bound	Conflict Rate per 1000 Opportunities (%) South-bound	Conflicts per Hour North-bound	Conflicts per Hour South-bound
Before	19.1	12.4	10.48	6.9
After				
Loops based on 85th percentile speed	11.2	5.0	7.36	3.2
Northbound loop based on 99th percentile speed, southbound unchanged	6.9	5.8	3.53	3.76

Obviously, the Kentucky approach is a compromise solution. Trial-and-error field work has produced the procedures now in use. It is important that additional research be conducted to determine truck dilemma zones by vehicle classification. The effect of excess grade as it affects the dilemma zone for all vehicles and the effects of short greens should also be the subject of additional research.

The new configuration proposed by Parsonson offers some advantages over most existing systems, particularly types 1 through 4. The 26-m (85-ft) stopline loops with quadrupole design provide much better control of the departing vehicle queue. This is also a far superior design for two-phase intersections with unprotected left-turn movements. To operate efficiently, it is essential that the stop-bar presence loops be deactivated once the waiting queue attains operating speed. Experience in Kentucky has shown that using the presence loop for extension and call will more likely result in a maximum time termination of the green phase rather than a gap termination. This is particularly true on a high-volume approach (12 000 average daily traffic).

It is recommended that the initial vehicle be set at 10 to 15 s to eliminate the problem of the short green, particularly when the signal display is visible for a great distance. It is also desirable that the signal rest or dwell in the dilemma-zone green phase.

Kentucky used and abandoned the stop-bar loop before the advent of the quadrupole configuration and digital self-turning loops. Poor maintenance performance and the short-green problem caused this abandonment. The large loops would not adequately detect all vehicles, and the analog non-self-turning loop amplifiers tended to detune during temperature changes. This generally resulted in a locked-in call and a maximum time termination of the green phase. It appears that detector technology has now evolved to the point that the maintenance factor can be virtually ignored in the process of selecting a detection scheme. In view of these facts, the quadrupole configuration is highly recommended for greater flexibility and operating efficiency.

Delay detection is considered essential for side-street phases at all times and for the main phases during off-peak, low-volume times.

Parsonson's advance detection strategy would be excellent for most vehicles approaching isolated, high-speed intersections. However, approach speeds in excess of 80 km/h (50 mph) are not rare at most intersections. Very comprehensive data should be collected to determine the existence of higher speeds, truck stopping problems, or excessive grades. If conditions warrant, a supplemental loop or loops should be considered. This loop extension time should be sufficient to allow a vehicle to pass the second loop using 80 km/h as the travel speed. The only deterrent to adding supplemental loops to the proposed configuration is the increased likelihood of maximum time termination of the green phase during periods of high traffic volume.

In conclusion, Parsonson has provided an excellent report that summarizes most known designs for dilemma-zone protection and indicates situations for which they would be most appropriate. It is anticipated that this report will assist in developing standardized designs. The new configuration proposed by Parsonson appears excellent and should provide excellent results when implemented.

REFERENCES

12. H. Sackman and others. Vehicle Detector Placement for High-Speed, Isolated, Traffic-Actuated Intersection Control: Volume 3—Case Study: Local Field Test and Evaluation. Federal Highway Administration, Rept. FHWA-RD-77-33, 92 pp., May 1977.
13. P. A. Winter and others. 1974 Brake Performance Levels for Trucks and Passenger Cars. Bureau of Motor Carrier Safety, Federal Highway Administration, U.S. Department of Transportation.

Author's Closure

Clark's discussion of high-speed truck traffic is an important contribution to the literature. He suggests that, where a signal is visible from 300 m (1000 ft) away, a minimum green time of 12 to 18 s is needed to meet the expectations of truck drivers. Heretofore, traffic engineering judgment seemed to center on 8 to 10 s as sufficient. Clark's discussion seems to be the profession's first perception that 8 to 10 s may not be enough in certain situations.

My paper proposes a new configuration intended for speeds of up to 80 km/h (50 mph). Clark points out that higher speeds need to be anticipated by the designer, particularly if trucks or downgrades are a factor. In response, I have modified my proposed configuration to that shown in Figure 12.

The upstream detector has been relocated to 115 m (380 ft) from the intersection, a distance adequate for vehicles approaching at 88.5 km/h (55 mph). The second detector is placed 77 m (254 ft) from the intersection; this is the upstream boundary of the dilemma zone for vehicles approaching at 56 km/h (35 mph). Both of these detectors are of normal design (i.e., not extended-call), and the unit extension of the (digital) controller is set at 1.9 s. It is easy to show that the 1.9-s extensions of the green will carry vehicles approaching at 64 km/h (40 mph) to 88.5 km/h through their respective dilemma zones. The allowable gap of 3.7 s for an 88.5-km/h stream and 3.9 s at 80 km/h (50 mph) is snappy enough to minimize the extension of green to the maximum interval.

The modified configuration could retain the 26-m (85-ft) long stop-bar loop proposed. However, the high cost and questionable durability of so long a detection loop are of concern. As an alternative, I propose a stop-bar loop only 8 to 9 m (25 to 30 ft) in length to be used with a novel hybrid detector. A loop of this length will usually bridge the gap between standing vehicles, ensuring a call. The detector is an extended-call and delayed-call (EC-DC) design with an adjustable timer for each of the two modes. As a queue discharges over the EC-DC stop-bar loop, the detector functions as an extended-call model. The stretch settings are high enough to produce an unbroken call until the vehicles are up to speed, at which point the detector gaps out. On gap-out, the detector becomes a delayed-call unit; the full-speed vehicles do not produce a call, and the detector is in effect disconnected.

In the proposed design, once the minimum green of 15 s has expired, the extended-call feature of the detector will hold in a call to the controller until there is adequate motion over the upstream detectors. Premature gap-out is avoided. Then, the EC-DC stop-bar detector will gap out, leaving only the upstream detectors to give dilemma-zone protection and control the allowable gap. The amount of stretch on the EC-DC stop-bar detector must be high enough to prevent premature gap-out but low enough to ensure that this detector will disconnect before interfering with the assignments of the upstream detectors.

There seems to be no evidence that such a hybrid detector has ever been built. During the spring of 1978, one was to be created by the traffic engineering staff of Gwinnett County, Georgia, in cooperation with the Canoga Controls Corporation. The effectiveness of the proposed design was then to be tested at a Gwinnett County intersection.

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

Relation Between Lighting Parameters and Transportation Performance

D. A. Schreuder, Netherlands Institute for Road Safety Research, Voorburg

The relation between the technical requirements for road-traffic lighting (geometry and photometry) and the functions of safety, speed, comfort, and cost is examined. Emphasis is placed on safety considerations. The "chain" between cost and effectiveness (i.e., transportation performance or accident reduction) is broken down into its elements. Each element can be studied separately, and the chain can be followed from both sides—supply and demand. On the supply side, cost leads to the conspicuity level provided; on the demand side, accident reduction leads to the conspicuity level required. Future recommendations must ensure that the level provided always exceeds the level required. The functional approach presented here promises results. It contrasts with the traditional approach, which considers only the visibility for standard tasks defined a priori—tasks that have no demonstrable relation to the driving task in real traffic situations. It is concluded that further detailed research is required.

Road lighting is expensive in terms of both money and energy. Therefore, these costs should be justified by benefits. Road lighting is thus considered utilitarian. Its benefits are found in four slightly overlapping areas: (a) road traffic and transportation performance, (b) public safety, (c) amenity, and (d) aesthetics. This paper is restricted to road traffic and transportation performance.

In the past, because road transportation was viewed from the economic viewpoint alone, cost-effectiveness considerations were simply a matter of bookkeeping. Recently, however, it has been realized that road transportation has an extremely wide impact on the community. The function of various facilities, such as road lighting, is to ensure that such transportation can function optimally. The function is usually described as allowing the road user to reach his or her destination safely, speedily, comfortably, and at minimum cost. Thus, cost-effectiveness considerations, and those for road lighting, are more complicated than bookkeeping only. It is usually assumed that all road-lighting requirements for safety, speed, and comfort are similar, increasing in that order in respect to their severity. Thus, safety can be considered the basic aspect and the others as only increasing the load on the lighting.

The effectiveness of road-lighting installations compared with no lighting at all can be estimated on the basis of traffic accident studies. Usually they are of the before-and-after type: The number of accidents before the installation of road lighting is compared with the number of accidents after installation, and appropriate correction is made for variations in travel, weather, and other changing factors on the road. As a result of methodological restrictions, the number of investigations that yield valid data is relatively small, but they all suggest a reduction of some 30 percent in nighttime accidents attributable to lighting (1, 2, 3, 4, 5, 6, 7, 8). This holds for "good" lighting installations compared with very little lighting or no lighting at all. However, to find out how good lighting should be in order to be considered good in considerations of cost-effectiveness, this approach does not give useful results. The reasons for this are that (a) the change in lighting installations proves to be applicable in before-and-after studies in only a few cases, (b) the number of accidents is too small and their registration not accurate enough to permit a rigorous statistical treatment, and, probably most important, (c) the effectiveness of lighting seems to depend not only on the lighting level but also on the type of road and traffic. Therefore, for a more detailed study, more detailed methods must be applied so that lighting installations of different quality in terms of accident-reduction potential can be compared. Such a detailed study requires the subdivision of the problem into a set (a chain) of subproblems. This chain is shown in Figure 1; the separate elements of the chain are described in detail in this paper. Further study areas pertain to including driving comfort and transport aspects.

This approach is not a very recent development. However, the pioneer work of Dunbar (9), Smith (10), and Waldram (11) passed unnoted, and usually—if it was considered at all—the aim of road lighting was supposed to be to approximate daylight as closely as possible. The more recent functional approach aims at a more realistic view (12, 13). The fundamental work of Hopkinson on discomfort effects (14) should also be mentioned.

ANALYSIS OF THE PROBLEM

One of the following questions should be asked, depending on whether the problem is to design or to assess lighting installations:

1. Which are the requirements on lighting and installation parameters (what are the costs) of a lighting installation that ensures a certain effectiveness (to be expressed in terms of accident reduction)?
2. What is the effectiveness of a lighting installation that shows certain characteristics in relation to lighting and installation parameters (and thus costs)?

Clearly, these two questions indicate two approaches to the problem that can be described as related to demand and supply. Equally clearly, lighting installations can be qualified as adequate or good only if the supply equals or exceeds the demand.

There are many ways to improve nighttime traffic conditions. Road lighting centers on the fact that nearly all information needed for participating in traffic (as driver or pedestrian) is of visual origin. Therefore, it seems natural to use the visual information supplied and required as the main concept. Because in most cases visual information is related to the degree to which objects are conspicuous, it is suggested that the amount of visual information should be expressed in levels of conspicuity.

In this way, the assessment of cost-effectiveness is split into two main problem areas:

1. How is the supplied conspicuity level related to lighting and installation parameters?
2. How is the demanded (required) conspicuity level

Figure 1. Chain of subproblems linking costs to road traffic and transportation performance.

"supply"
- [a] costs
- [b] installation parameters
- [c] lighting parameters
- [d] supplied conspicuity level

"demand"
- [4] required conspicuity level
- [3] manoeuvre behaviour
- [2] accident reduction capacity
- [1] road traffic and transportation performance

[a] ⟵ [b] ⟶ [c] ⟶ [d] ⟶

⟵ [4] ⟵ [3] ⟵ [2] ⟵ [1]

related to travel performance (expressed in terms of accidents or accident reduction)?

These two problem areas are connected by the requirement for adequate lighting: The supply should be equal to or larger than the demand (Figure 1). (It is assumed above that the actual costs of a specific lighting installation, for which the installation parameters are known, can be calculated.)

RELATION BETWEEN LIGHTING AND CONSPICUITY

The relation between lighting and conspicuity is in fact the supply part of the total chain. This part can be subdivided into a number of separate steps. Installation parameters are used here as a starting point in view of the fact that actual costs can usually be assessed when the installation parameters are known (see b and a in Figure 1).

The installation parameters represent the actual lighting installation. They include geometry (spacing, mounting height, road width, overhang, and arrangement), lamp or lantern characteristics (I_{80}, I_{88}, luminous flux, large-scale integration), and road surface characteristics (such as q_0 and x_p; q_0, S_1, S_2; or other characteristics). In most cases, all data are available; they will usually be available even before the installation becomes reality.

When the installation parameters are known (or selected), it is possible on the basis of the systems and programs proposed by the International Commission on Illumination (CIE) (15) to perform the next step—the assessment of the lighting parameters. The general system has been worked out in detail by CIE (16). This requires complete information on the lighting distribution (I tables), the reflection properties of the road (R tables), and, of course, other data.

It has been argued that a lighting installation can be described by a number of lighting (or photometric) characteristics such as the average road surface luminance and its uniformity, the glare control mark, the threshold increment, and the visual guidance (16, section 2). To a certain extent, visibility and driving performance have been taken into account in setting up these characteristics. Therefore, it is to be expected that a further and more systematic consideration of these aspects requires adaption (extension or change) of these characteristics. Furthermore, dynamic aspects have not been fully taken into account.

Photometric characteristics represent an intermediate step between the installation and conspicuity. For this purpose, they can serve rather well although in essence they are not a homogeneous set: The threshold increment is exclusively a matter of visual performance, luminance and uniformity combine aspects of visual performance and visual comfort, the G mark is exclusively a matter of visual comfort (by definition), and visual-optical guidance is a matter of traffic performance combined with visual comfort. In the past, however, the criteria have been considered to a certain extent to have a basic function of their own. Apart from the theoretical shortcomings of this view and the rigidity in lighting engineering they sometimes provoke, the major drawback of this way of looking at the matter is that the criteria are usually considered as independent factors, each of which calls for its own minimum value. Thus, CIE recommendations state that, for a particular type of road, L_{av} should exceed 2 cd/m^2, the uniformity should be better than 0.4, G should exceed 6, and TI should be lower than 10 percent. A more fundamental approach allows for investigations of the following type: If G is 8 and the uniformity 0.6, is it allowed or possible or advisable to decrease the minimum for L_{av} to 1.8 or 1.5 or 1.0? Obviously, answers to such questions are important for practical lighting design (3, 17).

Lighting parameters are considered to describe the visual environment in adequate detail to assess visibility (the visual guidance and the G mark play no part in this). This statement, although plausible, requires further confirmation. How far the statement can be applied depends on the accuracy desired. As a first approximation, the average road surface luminance is sufficient for many types of problems since it usually approaches the level of adaptation fairly well. On the other hand, for the description of the visual environment in actual traffic conditions, the characteristics given above are not sufficient: Dynamic effects are not included, and glare for other light sources and the influence of the luminance of the surrounds of the roads (shoulders and sidewalks) on the adaptation level are not fully known. Thus, results from this approach can be applied only for a restricted group of traffic situations. This should be kept in mind when, for example, the findings for busy urban streets are to be applied on rural motorways. It is precisely to handle these hitherto unknown factors that the approach from the demand side is being developed.

When the visual environment is defined, visibility can be assessed directly on the basis of the system adopted by CIE (18). The validity of the approach has been assessed in many investigations (3, 19, 20, 21, 22, 23, 24, 25, 26). Although some discrepancies did show up, in general there seems to be good agreement between the actual measurements and the theoretical framework that is developed primarily on the basis of laboratory experiments.

However, this approach to arriving at a set of requirements for road-lighting installations that ensure a preselected degree of road traffic and transport performance has come to a complete dead end. Although visibility can be assessed to a very precise degree, the results are of no practical value.

Visibility can be assessed only for a distinct object. Furthermore, the appropriate definition of the concept of visibility implicitly includes aspects of the task of the observer. It is customary to make certain assumptions in these two respects (usually the object is taken as a small cube or something similar, and visibility is taken as equivalent to threshold perceptibility). The results of this exercise are inconclusive in relation to road safety because it is impossible to find out from the visibility and lighting studies whether the assumptions are in any way related to what is

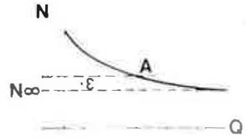


Figure 2. Qualitative indication of the relation between number of accidents N and quality of road lighting Q.

relevant in traffic. The only thing that emerges is the suspicion that visibility, defined in this way, has in fact very little to do with traffic.

Therefore, "field factors" of from 10 to 30 are included. These field factors actually reflect the common sense and the experience of the investigator both as a lighting designer and as a road user. This again explains why actual road-lighting installations usually perform quite well (as may be seen from the studies on accident statistics quoted earlier) although the fundamental questions were not answered at all. It also explains why important but rather precise questions like the minimum required levels of luminance for motorways (1 or 2 cd/m^2) cannot be answered and why new developments for which no experience exists can be perfected only by means of very expensive and time-consuming trials.

In summary, selecting (sets of) standard visual tasks a priori is useless, and selecting them on the basis of visibility considerations is dubious. The first does not give any information that can be applied with confidence in road situations; the other, representing in essence a circular argument, only serves to hide the real problems behind a curtain of beliefs and assumptions. The only valid basis for the definition of "standard" visual tasks is the actual requirement in road traffic.

The functional approach is a possible way out of this impasse. This will be a major part of future research in this field. In essence, it consists of considering the demand side of the conspicuity level (the term conspicuity level is preferred to the term visibility because one of the major problems at hand is to find aspects of visibility that are really relevant to road traffic).

RELATION BETWEEN CONSPICUITY AND ROAD SAFETY

As indicated earlier, of all the benefits of road lighting only road traffic and transportation performance are considered in this paper, which results in road safety being expressed as accident-reduction potential. A further restriction is now introduced: Drivers of vehicles (or automobiles) are considered to be users of road lighting—users meaning here those individuals that use the lighting to improve their possibilities for observation on the road. Thus, pedestrians are considered as objects and not as road users. All these restrictions are not of fundamental value; they are introduced only to reduce the size of the discussion. All arguments and all conclusions can be restated in such a way that they include other types of benefits, other criteria of quality of travel, and other road users, the common idea being the fact that in all cases the lighting serves a well-defined purpose and is therefore utilitarian.

The benefits of road lighting for automobile drivers can be expressed in the number of accidents that are prevented by the lighting. This has loosely been described as the accident-reduction potential of the lighting. More precisely, these benefits could be expressed as $N = f(Q)$ where N is the number of accidents still occurring and Q the quality of the lighting. The first research task is to define Q in such terms that it can be applied.

As indicated earlier, this functional relation cannot be established directly from accident statistics. Not only the description of Q is lacking; to approximate a function relation, the "steps" in Q, and thus the differences in consecutive steps in N, should be small. This holds even more if one looks for the minimum admissible value of Q. For this, the relation is usually taken as having an asymptote N_∞ in N for high Q. The minimum admissible value of Q is that value where $(N - N_\infty) < \epsilon$, ϵ being small and depending in magnitude on the amount of social concern. This implies that when $N = f(Q)$ is not known as a real (continuous) function, it should at least be known in steps smaller than ϵ. This is shown in Figure 2. The establishment of this (quasi-) function directly from accident statistics requires an enormous experimental effort. The reason for this is that accidents, being occurrences that happen relatively seldom, can be described by a stochastic process (a Poisson distribution that can be approximated by a normal distribution). To distinguish between two normal distributions that differ only slightly in their mean values (the step ϵ), the samples to be taken must be large. The length of road network available for the experiments is also large. This makes it virtually impossible to perform this analysis within reasonable time and cost limits.

As a possible way out of these difficulties, it is suggested that the relation between conspicuity and road safety be broken up into a number of separate steps, as shown in Figure 1. It is also suggested that the analysis of the driving task be included as one of the intermediate steps. As has been argued in other places in great detail (27, 28, 29, 30), the driving task can be described in the hierarchy of decision processes given below:

Individual Behavior	Collective Behavior
Selection of motive	Trip generation
Selection of destination	Trip distribution
Selection of mode of transport	Modal split
Selection of route	Assignment
Selection of maneuver	Traffic flow

The hierarchical level of most importance here is the lower one in which the maneuvers are described. Thus, the actual handling of a vehicle can be described as a series of maneuvers, each of which is performed after a decision to do so, a decision based among other things on (visual) information about the outside world.

The "space" required for the adequate performance of each maneuver can be defined, as can the available space. Space should be understood here in a very general sense; it is determined not only by the border of the roadway but also by the maneuverability of the vehicle, the ability of the driver, the presence and the maneuvers (actual or planned) of other road users, visiblity, meteorological conditions, the skidding resistance of the road surface, and other factors. The actual extent of both the required and the available space is unknown. The driver has to base the decision whether or not to undertake a certain maneuver on estimations of the extent of the space. It may be assumed that the estimation of the required space is not a matter of visibility but of confidence in the road-holding capability of the vehicle, the driving ability of the driver, and so on. The estimation of the available space, however, is clearly a matter of visibility. There are three possibilities: The actual extent A of the available space is larger than, smaller than, or equal to the estimated extent A′. A more detailed consideration leads to the preference for A′ = A. Thus, the road lighting should be such that A can be estimated correctly.

This idea is, in a vague way, behind the requirement that the visibility distance of objects must be at least equal to the minimum stopping distance. However, if one selects a visual task that corresponds to an object for which the driver really has to stop, e.g., a truck parked on the roadway, the visibility distance becomes unrealistically large. Furthermore, trucks have signal lights or at least reflectors. Therefore, one usually selects a very small object—e.g., the notorious 20- by 20-cm^2 dull grey box (not an object drivers usually have to stop for). This is precisely the impasse indicated earlier.

The way out is the consideration that there are many objects that can present themselves and that there are a number of possible maneuvers from which the driver has to select one after he or she has had the opportunity to see and recognize the object and has had the opportunity to make an assessment of the pros and cons of the different maneuvers. It is the analysis of the driving task that permits one to state which are the possible maneuvers under certain circumstances and which is the most appropriate. For the different maneuvers and for the different conditions under which they have (or may have) to be performed, the required space to maneuver can be assessed by taking into account the actual or the average value of vehicle performance, road characteristics, and driver ability. By taking into account the characteristics of the object that requires the particular maneuver, the visual environment can be described so as to enable the actual or the average driver to really observe the object. This visual environment corresponds with the demand side, with the required conspicuity level. Finally, the lighting installation should be such that demand does not exceed supply.

In this way, the chain that links installation parameters to road traffic and transportation performance is complete. It should be noted, however, that in the analysis given above specific aspects of vision and lighting are involved only in the last two steps. A major part of the future research mentioned in this paper is on the schedule of the CIE Technical Committee on Road Lighting.

CONCLUSIONS

The effectiveness of good road lighting compared with no lighting at all can be determined from traffic accident studies. When one seeks to know how good is good in this respect, accidents are not frequent enough and not recorded accurately enough.

The chain between costs and transportation performance is split up into smaller parts; the costs (and the installation parameters closely related to them) stand for the supply side, and road traffic and transportation performance (and accident-reduction potential, which is closely related) form the demand side. For good road lighting, the supply should equal or exceed the demand. The chain is followed by starting from both sides simultaneously. The supply side gives the supplied conspicuity level; the demand side leads to the required conspicuity level. Again, the supply should equal or exceed the demand.

The supplied conspicuity level can be derived from installation parameters by means of well-established methods. A similar derivation of the required conspicuity level from traffic and transportation performance requires further research. The traditional method, in which one or maybe two standard visual tasks are postulated as being representative for driving, is completely unsatisfactory and may even lead to erroneous results.

REFERENCES

1. Street Lighting and Accidents. International Commission on Illumination, Paris, Publ. 8, 1960.
2. P. R. Cornwell and G. M. Mackay. Lighting and Road Traffic I. Traffic Engineering and Control, 1972, pp. 142-144.
3. J. B. DeBoer, ed. Public Lighting. Centrex, 1967, Eindhoven, the Netherlands.
4. P. Box. Relationship Between Illumination and Freeway Accidents. Illuminating Engineering, Vol. 61, 1971, pp. 365-393.
5. Lighting, Visibility and Accidents. Organization for Economic Cooperation and Development, Paris, 1971.
6. J. B. DeBoer and D. A. Schreuder. New Developments in Road and Street Lighting With a View to Road Safety. In Intertraffic '72 RAI, Amsterdam, 1972.
7. T. H. Tan. Wegverlichting in de Praktijk. In Wegon twerp en Wegverlichting Tegen de Achtergrond van de Verkeersveiligheid, het Nederlandsche Wegencongres, The Hague, 1974.
8. A. Fisher. Road Lighting as an Accident Countermeasure. In Measures of Road Lighting Effectiveness, 1977 Symposium, Karlsruhe, West Germany, International Commission on Illumination, Paris, 1977.
9. C. Dunbar. Necessary Values of Brightness Contrast in Artificially Lighted Streets. Trans., Illuminating Engineering Society, London, Vol. 3, 1938, p. 187.
10. F. C. Smith. Reflection Factors and Revealing Power. Trans., Illuminating Engineering Society, London, Vol. 3, 1938, pp. 196-200.
11. J. M. Waldram. The Revealing Power of Street Lighting Installations. Trans., Illuminating Engineering Society, London, Vol. 3, 1938, p. 173.
12. D. A. Schreuder. Road Lighting and Traffic Safety—A Functional Approach. Lux 57, 1970, pp. 146-147 and 256-263.
13. D. A. Schreuder. A Functional Approach to Lighting Research. Tenth International Study Week in Traffic and Safety Engineering, World Touring and Automobile Organization, Rotterdam, 1970.
14. R. G. Hopkinson. Discomfort Glare in Lighted Streets. Trans., Illuminating Engineering Society, London, Vol. 5, 1940, pp. 1-24.
15. Recommendations for the Lighting of Roads for Motorized Traffic. International Commission on Illumination, Paris, Publ. 12-2, 1977.
16. Calculation and Measurement of Luminance and Illuminance in Road Lighting. International Commission on Illumination, Paris, Publ. 30, 1976.
17. E. Enzmann and W. Adrian. The Influence of Glare on Visual Performance and Lighting Parameters on Roads of Equal Luminance Uniformity. In Measures of Road Lighting Effectiveness, 1977 Symposium, Karlsruhe, West Germany, International Commission on Illumination, Paris, 1977.
18. A Unified Framework of Methods for Evaluating Visual Performance Aspects of Lighting. International Commission on Illumination, Paris, Publ. 19, 1972.
19. J. B. DeBoer. Fundamental Experiments on Visibility and Admissible Glare in Road Lighting. International Commission on Illumination, Stockholm, 1951.
20. J. B. DeBoer and others. Appraisal of the Quality

of Public Lighting Based on Road Surface Luminance and Glare. International Commission on Illumination, Brussels, 1959.
21. R. Walthert. Zur Bewertung der Leuchtdichteverteilung Beleuchteter Strassen. Univ. of Karlsruhe, thesis, 1973.
22. O. M. Blackwell and H. R. Blackwell. A Proposed Procedure for Predicting Performance Aspects of Roadway Lighting in Terms of Visibility. Journal of Illuminating Engineering Society, London, 1977, pp. 148-166.
23. H. R. Blackwell and O. M. Blackwell. A Basic Task Performance Assessment of Roadway Luminous Environment. In Measures of Road Lighting Effectiveness, 1977 Symposium, Karlsruhe, West Germany, International Commission on Illumination, Paris, 1977.
24. V. Gallagher and P. G. Meguire. Contrast Requirements of Urban Driving. TRB, Special Rept. 156, 1975, pp. 40-52.
25. A. Ketvirtis and P. J. Cooper. Detection of Critical-Size Object as Criterion for Determining Driver's Visual Needs. Presented at 56th Annual Meeting, TRB, Jan. 24-28, 1977.
26. I. Economopoulos. Relationship Between Lighting Parameters and Visual Performance in Road Lighting. In Measures of Road Lighting Effectiveness, 1977 Symposium, Karlsruhe, West Germany, International Commission on Illumination, Paris, 1977.
27. E. Asmussen. Beleid Onderbouwd. Proc., Congress on Future in Road Safety, Netherlands Institute for Road Safety Research, Voorburg, Publ. 1976-4N, 1976.
28. D. J. Griep. Analysis of the Driving Task: System-Analytical Points of View. Presented to Organization for Economic Cooperation and Development Symposium on Road User Perception and Decision Making, Rome, Nov. 13-15, 1972, Netherlands Institute for Road Safety Research, Voorburg, 1972.
29. D. A. Schreuder. De Rol van Functionele Eisen Bij de Wegverlichting. In Wegontwerp en Wegverlichting Tegen de Achtergrond van de Verkeersveiligheid, het Nederlandsche Wegencongres, The Hague, 1974.
30. D. A. Schreuder. Signaleren in het Wegverkeer. Electrotechniek, Vol. 55, 1977, pp. 254-275.

Publication of this paper sponsored by Committee on Visibility.

Abridgment

Reanalysis of California Driver-Vision Data: General Findings

Brian L. Hills, Transport and Road Research Laboratory, Crowthorne, England
Albert Burg, University of California, Los Angeles

Early studies of the relations between driver vision and accidents were contradictory in their findings, largely because of the small sample sizes used. However, in 1967 and 1968, Burg (1, 2) published the findings of a study in which visual measurements made on over 17 500 California drivers were compared with their 3-year driving records, which included over 5200 accidents. It remains to a considerable extent the most comprehensive study of driver vision yet accomplished.

Taking the driving population as a whole, Burg found very weak but statistically significant correlations between various vision scores and driving records. The vision test that best predicted accidents proved to be a nonstandard one—dynamic visual acuity (DVA), in which the observer had to resolve detail in a rapidly moving acuity target; however, by itself, DVA remained a poor predictor of a driver's accident rate. This and other general findings of Burg's study reflected both the multicausal nature of traffic accidents and the need to develop tests of visual perception that are more relevant to the driving task than the classical tests of vision (which were largely devised for reading purposes).

Vision standards for driver licensing require not only the selection of valid visual characteristics to be tested but also the establishment of valid cutoff scores as criteria for passing or failing. To date there has been virtually no research into the latter problem, and this study was conducted with this need in mind. This paper summarizes the major findings of the study and is taken from a more detailed report (3).

The study explored in depth the implications of Burg's data for driver-vision standards and concentrated on determining whether certain subgroups of the driving population displayed stronger relations between vision and driving than did others. Preliminary work suggested that analysis of older drivers rather than of those with poor vision was most likely to show these stronger relations. Therefore, in the main analyses, the sample was divided into four age groups: under age 25, ages 25 to 39, ages 40 to 54, and over age 54.

VISION TESTS

The vision tests used by Burg included the following:

1. Static visual acuity (SVA)—binocular distance acuity measured by using the Bausch and Lomb Ortho-Rater (F-3 test) and, for a subsample of the total population, a Snellen chart;
2. Dynamic visual acuity (DVA)—the ability to perceive a series of rapidly moving (Ortho-Rater) checkerboard acuity targets projected on a cylindrical screen at two angular rotation speeds: 90°/s and 120°/s;
3. Low-light recognition threshold—the threshold amount of light required to recognize familiar targets;
4. Glare recovery—the length of time taken by the subject to reattain the low-light recognition threshold after exposure to 5-s glare; and

5. Field of vision—total horizontal visual field, both eyes combined.

DRIVING RECORD VARIABLES

Research has clearly shown that the number of kilometers driven (quantitative exposure to risk) is a fundamental factor in predicting the number of accidents and convictions for traffic violations experienced by a driver. In this study, therefore, only accident and conviction rates [per 161 300 km (100 000 miles) driven] were considered as driving record variables, to minimize the influence of exposure. Furthermore, 98 drivers with average annual distances driven of less than 1613 km (1000 miles) were excluded from the analyses since they were found to be highly atypical of the total sample on a number of critical driving and personal characteristics. The total sample ultimately analyzed numbered 14 283.

COMPARISON OF STATIC ACUITY TESTS

The primary test instrument for determining binocular static acuity in the Burg study was the Bausch and Lomb Ortho-Rater, which uses checkerboard visual targets (the same instrument was used for official testing of applicants for drivers' licenses in California where the study was conducted). To permit a comparison of two common test methods, binocular static acuity was also measured for a subsample of 4753 subjects by using a Snellen chart. In view of the fact that current proposals for the harmonization of European standards are based on wall charts that use either the Snellen alphabetic characters or Landolt C-rings, the results of the comparison are of interest.

As might be expected, the results showed that there is an approximately linear relation between the two acuities: however, the product-moment correlation coefficient is only 0.70, hardly enough for accurate prediction of one score from another. Although the two acuities coincide at 20/20, the slope is not 45°; 20/40 Snellen was found to be approximately equivalent to 20/30 Ortho-Rater, whereas 20/40 Ortho-Rater was equivalent to about 20/67 Snellen. While 46 of the sample would have failed a standard of 20/40 Snellen, only 18 would have failed a 20/40 Ortho-Rater standard, a dramatic difference in failure rate. These differences are believed to be attributable to the repeated pattern of the checkerboard, which would appear to be more resistant to spherical or certain cylindrical blurrings than the more complex form of alphabetic letters.

This divergence between Snellen and Ortho-Rater acuities at the poor-vision end of the spectrum is clearly of significance when international comparisons are made of standards for driver vision and must be taken into account in considering the analyses for SVA that follow. It is not known whether this relation would also apply to different measures of dynamic acuity.

DATA ANALYSES

The primary data analyses investigated relations between vision and driving as a function of both age and level of visual performance and used correlational analysis and t-tests to determine the statistical significance of differences in mean accident rates for various subgroups of the sample. The results were quite different for the over 54 age group compared with those for the three younger age groups; therefore, this age group is discussed separately.

Drivers Age 54 and Younger

The data analyses revealed no significant relation between accident rate and any of the visual performance measures studied for the three younger age groups. It is felt that for young drivers factors other than vision, such as experience, are likely to be more highly correlated with accidents whereas for all age groups any deterioration in visual performance might be at least partially compensated for by modifications in driving behavior (such as reduced speed and increased headways), by changes in looking behavior, and perhaps by improved manipulative skills. It may well prove that a higher order visual test—for example, a test of hazard perception—is a more effective accident predictor for these younger age groups than the tests of more basic visual abilities examined in this study (4). Henderson and Burg (5) suggest that tests of perceptual ability rather than sensory capability are more likely to be related to driving performance, which suggests the need for more complex performance tests that involve cognitive as well as sensory aspects.

Drivers Over Age 54

Weak relations were found between certain of the visual tests and accident rates for the oldest age group. DVA and SVA tests showed the most systematic and consistent relations; 90°/s DVA exhibited a slight superiority. Although they were significant, the correlation coefficients found were very low, which indicates that for an individual driver the accident prediction value of these tests is poor. A more detailed age analysis failed to define more precisely the age at which these relations develop although evidence was found to suggest that there are marked differences in the way they develop under daytime and nighttime conditions.

The results for the two tests of night vision—low-light recognition threshold and glare recovery—were regarded as inconclusive for the over 54 age group although glare recovery was the more promising of the tests. The test of total visual field is discussed below.

Total Visual Field

For the over 54 age group, there was no evidence of a progressive increase in accident rate with decreasing total visual field. In addition, no evidence was found to support a vision standard of 140° (as adopted by a number of states and recommended by the World Health Organization, the American Optometric Association, and a number of other bodies). These findings are in general agreement with those obtained by previous researchers.

Developing Cutoff Scores for Driver Licensing

A systematic attempt was made to find out whether vision test cutoff scores that might be considered valid for purposes of driver screening could be determined. This was done by systematically varying the pass score from the highest to the lowest levels of visual performance for each test and then using t-tests to determine for each pass score the statistical significance of the difference in mean accident rate between the pass and fail groups.

The results of these analyses were in keeping with those of the correlational analyses described above. For the three younger age groups, the vision tests provided no cutoff scores that could be considered consistently valid and useful. For the over 54 age group, however, useful cutoff scores were found for both static and dynamic acuity. The most consistent results were obtained for 90°/s DVA, where both 20/50 and 20/67 cutoff scores proved highly significant. The 20/67 score failed 6 percent of the age group, and the mean accident rate for this fail group was twice that of the pass group.

An Ortho-Rater static acuity cutoff score of 20/40 placed 1.6 percent of the drivers over age 54 in the fail group, and their accident rate was 2.5 times that of the pass group. But since it was determined that this result was applicable solely to male drivers, its usefulness for female drivers is questionable.

For total visual field, a cutoff score of 170° was highly significant; however, such a standard would fail nearly 80 percent of the over 54 age group, an obviously impractical ratio of selection. The results for low-light recognition threshold and glare recovery were not consistent. -

DISCUSSION OF RESULTS

If a vision test is to be successful in screening applicants for drivers' licenses, it must correctly identify a maximum number of drivers in the target group, i.e., those with both bad vision and unacceptable accident rates, while at the same time minimizing the number of drivers identified as having bad vision but acceptable accident rates. Table 1 illustrates this point by giving the number and percentage of drivers over age 54 who fall into each of several categories based on combinations of vision score and accident rate. Four values are chosen as examples of acceptable accident rate: the sample mean and two, three, and four times the sample mean. A score of 20/40 Ortho-Rater static acuity is used as the vision test cutoff (passing) score because it was found to be highly significant and gave the largest difference in mean accident rate (2.5:1) between the fail and pass groups of all the static acuity levels tested. (However, it is possible that more significant results might have been obtained at a level of acuity intermediate to those available in the Ortho-Rater.)

Data given in Table 1 show that using an accident rate of 3.88 to define the maximum acceptable and then the 20/40 cutoff identifies only 8 of the 250 drivers with unacceptable accident rates and would reject 40 drivers with acceptable rates. A simple index of merit is shown that suggests that this is the best performance of the test for the four definitions of acceptable accident rate considered. However, such a simple index is not likely to be an adequate one in view of the social costs of denying a license to 40 acceptable drivers in order to remove 8 unacceptable drivers from the road. A more valid index must take into account all of the social and political costs and benefits associated with each category.

As indicated earlier, accident rate rather than accident frequency was used as the accident-record criterion because the former takes into account exposure to risk (distance driven) as a factor causing accidents. The analyses supported this decision by demonstrating that, if accident frequency is used as the criterion instead of accident rate, then the 20/40 Ortho-Rater cutoff score is even less successful in identifying drivers with unacceptable accident experience. For example, at the expense of denying a license to 47 acceptable drivers, this cutoff score identifies only 1 of the 86 drivers over age 54 who had more than one accident in 3 years.

It should be pointed out that older drivers drive much less than do younger drivers (1), and the use of accident rates as a basis for vision standards can therefore lead to the paradoxical situation in which older drivers who failed a test would have fewer accidents per year than younger drivers who passed it. This raises a number of social and political issues that are outside the scope of this study.

SUMMARY

In summary, it must be said that, as a basis for vision standards that are valid in terms of potential accidents saved, the tests studied must be regarded as disappointing. The failure to find a direct relation between poor visual performance and high accident rate for young and middle-aged drivers has been consistent throughout the study, and, for the over 54 age group, the relations obtained are significant but weak. The ability of these tests to identify drivers likely to have accidents—without paying an unacceptably high penalty in the rejection of good drivers—remains questionable.

These findings lend support to current attempts to find perceptual tests of visual performance that are much better accident predictors than the largely classical sensory tests of vision studied here. (Tests of contrast sensitivity, movement perception, and hazard perception are among those currently being examined, and it is recommended that investigation of other stimulus conditions for the promising glare-recovery test be carried out.)

It should be stressed that the significant relations found for older drivers may not be causal. A driver's visual performance in this age range may merely reflect his or her "effective" (or phenomenal) age, and some other factor such as deterioration of the brain's central processing capacity may be the fundamental cause of increased accident rates. Thus, improving a driver's visual performance may not improve his or her accident rate; however, even if a measure of visual performance is not causally related to accident rate, any predictive power it may have could still be valuable for the purposes of screening or visual standards.

Table 1. Effectiveness of Ortho-Rater binocular SVA cutoff score in differentiating drivers over age 54 with acceptable and unacceptable accident rates.

Acceptable Accident Rate (accidents 161 300 km)	Over 54 Age Group								Simple Merit Index[b]
	Pass Score (20 40 or better)				Fail Score (worse than 20 40)				
	a (drivers with acceptable accident rate)		b (drivers with unacceptable accident rate)		c (drivers with acceptable accident rate)		d (drivers with unacceptable accident rate)[a]		
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	
≤0.97 (sample mean)	2436	81.9	491	16.5	38	1.28	10	0.34	1.31
≤1.94	2501	84.1	426	14.3	38	1.28	10	0.34	1.55
≤2.91	2607	87.6	320	10.8	40	1.34	8	0.27	1.63
≤3.88	2685	90.3	242	8.13	40	1.34	8	0.27	2.22

Note: 1 km = 0.62 mile.

[a]Target group. [b][(d/c)/(b/a)].

ACKNOWLEDGMENTS

The data analyzed in this study were originally collected under a grant from the U.S. Public Health Service and a contract with the state of California and the U.S. Bureau of Public Roads. The work described in this paper forms part of the program of the Transport and Road Research Laboratory, and the paper is published by permission of the laboratory director.

REFERENCES

1. A. Burg. The Relationship Between Vision Test Scores and Driving Record: General Findings. Department of Engineering, Univ. of California, Los Angeles, Rept. 67-24, 1967.
2. A. Burg. Vision Test Scores and Driving Record: Additional Findings. Department of Engineering, Univ. of California, Los Angeles, Rept. 68-27, 1968.
3. B. L. Hills and A. Burg. A Reanalysis of California Driver Vision Data: General Findings. Transport and Road Research Laboratory, Crowthorne, England, Rept. 768, 1977.
4. B. L. Hills. Some Studies of Movement Perception, Age and Accidents. Transport and Road Research Laboratory, Crowthorne, England, Rept. SR 137UC, 1975.
5. R. L. Henderson and A. Burg. Vision and Audition in Driving. System Development Corp., Santa Monica, CA, Rept. TM-(L)-5297/000/00, 1974.

Publication of this paper sponsored by Committee on Visibility.

Abridgment

Roadside Hazard Model

Jerome W. Hall, University of New Mexico
Thomas E. Mulinazzi, University of Maryland

One category of traffic accidents that has received increased attention in recent years is the collision of a single vehicle with an object adjacent to the roadway. These single-vehicle, fixed-object (SVFO) accidents constitute approximately 17 percent of all reported accidents, and the probability of occupant injury in these accidents is significantly higher than is the probability for the complementary set of accidents. In an effort to develop cost-effective solutions to this problem, the Maryland Department of Transportation sponsored a study of these collisions on state-administered roads other than freeways. The objective of the study was to identify and quantify the parameters associated with SVFO accident severity and probability and to incorporate them into a hazard model. Previous reports (4, 5) have described the preliminary findings, and this abridgment presents the results of the concluding phase of the study.

INPUTS TO A ROADSIDE HAZARD MODEL

Field surveys conducted as part of the first phase of this study identified numerous objects adjacent to the roadway. A majority of these objects, including drainage facilities, traffic signal supports, and utility poles, were manmade. The number of these elements, coupled with the cost and logistical problems of their removal, relocation, or redesign, requires that attention be devoted to those elements that (a) result in injury to the occupants of striking vehicles and (b) are relatively more likely to be struck.

Severity

The degree to which a particular type of object results in injury to vehicle occupants can be quantified by its severity index (SI). From 1970 to 1975, reported SVFO accidents on Maryland and U.S. routes had an average SI of 0.44. The severity indexes determined from accident records are average values for all reported SVFO accidents. Caution must be exercised in using these averages primarily because of a significant number of unreported accidents.

All other factors being equal, accidents at higher speed will result in a larger frequency of injuries. Rural highways have more severe accidents although some SVFO accidents on 47- to 56-km/h suburban arterials, especially those that occur at night when traffic volumes are relatively low and involve drivers who are in "other than normal" condition, occur at high speeds. Accident records indicate that 44 percent of SVFO accidents involve drivers who are traveling at speeds too fast for conditions. A general model for determining the priority of roadside-hazard improvements must incorporate some speed-related parameter to highlight locations where SVFO accidents are likely to be more severe.

The most serious problem that is not reflected in accident records or accounted for by the SI is the variation in object design. For example, a variety of guardrail designs are used; W-beam designs are the most common, but single- and multiple-wire cable guardrails are also used. Various mounting heights are used in conjunction with blunt, flared, or buried terminals. Similar variations exist for the designs of other fixed objects and are of considerable importance because they affect the severity of SVFO accidents.

Probability of Impact

It is also essential for the hazard model to incorporate the likelihood of impact with a fixed object. Based on this research, the most important factors are traffic exposure, roadway geometrics, and placement of fixed objects.

The extent to which traffic is exposed to the object is partially reflected by the traffic volume on the route. However, volume by itself is not directly related to SVFO accident experience since multiple-vehicle accident experience increases at higher volumes whereas single-vehicle accidents decrease. Traffic volume is also related to roadway characteristics—notably road width and shoulders—that are associated with the frequency of roadside encroachments (3, 7). This research found that an unusually high percentage of SVFO accidents (62 percent) occur during conditions other than daylight. On some study sections, 80 percent of SVFO accidents occur during hours of darkness.

Studies that have concentrated on accidents and the geometrics of rural highways (1) have found that alignment and roadway width are the most significant factors. The field investigations in this study found that the adverse features of roadway alignment—notably steeper downgrades, sharp horizontal curvature, and the absence of shoulders—are the most critical factors.

Placement of objects involves three components that influence the probability of impact and warrant inclusion in the roadside hazard model: (a) the distance of the object from the edge of the traveled way, (b) the placement of the object inside versus outside a curve, and (c) the presence or absence of curbs or guardrail protecting the object.

MODEL DESCRIPTION

Results of previous studies (2) prompted the following conclusions with respect to the SVFO relative hazard model:

1. Recognition must be given to the probability and severity of impact;
2. It is essential to minimize the data items that must be collected for each object while maintaining the accuracy of the model; and
3. Because of problems with the reported frequency of SVFO accidents, verification of the model will be difficult.

The initial structure for the roadside hazard model is

$$H = K f_1(D) f_2(S) f_3(SI) f_4(V) f_5(G) \qquad (1)$$

where

H = relative hazard of a particular object,
K = a normalizing constant,
D = distance of the object from the road edge,
S = prevailing speed of traffic on the roadway,
SI = severity index associated with the type of object,
V = volume of traffic, and
G = geometric conditions.

Quantification of Parameters

In determining the values of the factors to be used in the model, the following considerations are of prime importance:

1. Each factor must be based on data that can be easily obtained from field studies and the existing record system.
2. For a given parameter, the factors must recognize in a logical manner the varying level of hazard associated with that parameter.
3. The quantification must recognize that individual parameters are not necessarily independent nor of equal importance.
4. The resultant hazard index can be normalized but should be proportional to the combined effect of the expected frequency and severity of accidents.

Distance

An object close to the roadway is more hazardous than a similar one that is farther removed. The relevant distance is measured from the right-hand edge of the travel lane to the object's nearest point. Since exact measurement for each object would be time-consuming and would not increase reliability in proportion to the effort involved, it is recommended that distance ranges be used. An analysis of distance-exceedance distributions provided the basis for quantifying the distance factor, as given below:

D (m)	f_1(D)
< 1.5	1.00
1.5-3.0	0.76
3.0-9.0	0.33
> 9.0	0.12

Speed

The factor of speed is important to the roadside hazard model because it affects the time an errant driver has to perceive and react and is related to the kinetic energy dissipated by a collision. Because of the limited data available, the posted speed limit, which is a reasonable representation of speeds on most state highways, was used in the model rather than the distribution of speeds of vehicles leaving the roadway. Since the speed factor is primarily intended to reflect severity and secondarily to account for probability of impact, the inclusion of these two considerations is achieved by using the parameter $(S + 16)^2$ where S is the posted speed limit. The rationale for this parameter is the reported higher accident experience at speeds 16 km/h faster than the posted speed limit. Using an assumed maximum speed of 80 km/h gives the following values of this parameter:

S (km/h)	f_2(S)	S (km/h)	f_2(S)
48	0.44	72	0.84
56	0.56	80	1.00
64	0.69	88	1.17

Severity Index

The SI for reported SVFO accidents serves as the best criterion for judging the seriousness of accidents that involve the various types of objects. It can be readily obtained from the accident-record system and can be periodically updated as new data become available. Using data for 20 000 SVFO accidents and the SI of 0.55 for light supports as the normalizing value gives the following calculated values of f_3(SI):

SI	Type of Object	f_3(SI)
0.271	Construction barrier	0.49
0.280	Other fixed object	0.51
0.283	Sign support	0.52
0.309	Fence	0.56
0.353	Curb or wall	0.64
0.379	Building	0.69
0.399	Guardrail	0.73
0.463	Culvert or ditch	0.84
0.506	Embankment	0.92
0.513	Bridge	0.93
0.529	Other poles	0.96
0.533	Tree or shrubbery	0.97
0.550	Light support	1.00

Traffic Volume

Traffic volume is included in the roadside hazard model because it is related to the rate of encroachment (although the latter is exceedingly difficult to measure) (6). This research has found that 52 percent of all SVFO accidents (versus 20 percent of all other accidents) occur between 9:00 p.m. and 7:00 a.m. Although reliability is improved by incorporating nighttime traffic volumes, the simplest procedure would employ only type of roadway and estimated average daily traffic (ADT). The volume factors given below were determined from an analysis of SVFO accident rates and normalized to a base of 25 000 ADT:

Type of Roadway	Adjustment Factor
Multilane	0.040 x (ADT in 000s)
Wide rural	0.064 x (ADT in 000s)
Narrow rural	0.088 x (ADT in 000s)

Geometrics

The principal geometric conditions of the roadway related to SVFO accident experience are roadway alignment and the placement of the fixed object. Data from this study have been combined with findings reported by Wright (9) to assess the relative hazard of these various conditions. The following matrix gives $f_5(G)$ as a function of roadway grade and curvature and placement of the fixed object:

		Grade (%)		
Curvature	Placement	< -2	-2 to -5	> -5
0°-3°	Inside	0.108	0.135	0.215
	Tangent	0.133	0.167	0.265
	Outside	0.250	0.315	0.500
3°-6°	Inside	0.129	0.163	0.258
	Tangent	0.159	0.200	0.318
	Outside	0.300	0.378	0.600
> 6°	Inside	0.215	0.271	0.430
	Tangent	0.265	0.334	0.530
	Outside	0.500	0.630	1.000

Other Parameters

The most obvious factor not directly accounted for in the model is the distinction between spot and continuous objects. The study found that 42 percent of the SVFO accidents involved spot objects. In comparison with freeways, the distinction loses significance because some suburban roadway sections had more than 190 fixed objects/one-directional km, and rural sections had up to 60 objects/one-directional km. A second parameter not adequately addressed by the model is differences in the design of the fixed object. For example, the model does not indicate a reduction in hazard if wire guardrail is replaced by a more modern installation. A third element that is not considered at this stage in the model is the relative hazard of objects placed on the foreslope versus the backslope. The latter is intuitively a better condition, but this research was unable to quantify the difference. These shortcomings are all accommodated to some extent in other models designed for limited-access facilities (8).

USE OF THE MODEL

The hazard rating has three basic uses. Of primary interest is the fact that it can use field data to determine the relative hazard associated with the various fixed objects along the roadside, thus establishing a priority ranking for improvement. Second, the model permits a relative assessment of the various forms of remedial action, including the effects of severity or accident reduction. Third, the model can be applied to a variety of roadway design and operating features to develop a hazard hierarchy for fixed objects.

Field Data Collection

In the development of the model, major emphasis was placed on minimizing field data collection while maintaining reliability. The data needed include route characteristics (speed limit and traffic volumes), type and placement of objects, and geometric design features. The field data are recorded on a suitably designed form by a two-member survey crew who travel the roadway in a properly instrumented vehicle. Essential equipment includes an accurate odometer, a slope meter, and equipment for measuring lateral distance.

Despite efforts to simplify the data requirements of the model, a substantial amount of information will have to be collected, especially on roadway sections that have large numbers of fixed objects within 9 m of the roadway. On several of the study routes, there was less than 4.5 m of right-of-way adjacent to the pavement. This consideration, coupled with a hazard model analysis, led to the recommendation that initial data collection efforts be limited to objects that are within the existing highway right-of-way, or 4.5 m, whichever is less. A second limitation to facilitate data collection is the adoption of a policy for the correction of easily identifiable objects that use hazardous designs (e.g., deficient guardrail). A third possibility for expediting the inventory would be an automated field data collection system that would directly create a file for computer processing.

Application

The model can be applied on a theoretical basis to determine the effect of various forms of remedial action and to establish a ranking of relative hazard. The specific inputs in this analysis are the calculated hazard index reduction and considerations of practicality and economics. Although specific characteristics at a particular location may dictate otherwise, a general cost-effective structure for remedial action that is in general agreement with published guidelines for fixed-object correction was developed (5).

Interaction of Model Parameters

The application of the model provides a method for obtaining some insight into which combinations of fixed objects and other parameters warrant the most immediate attention. To use the model in this manner, a variety of roadway-volume classifications were considered. For each category of roadway volume, there are 8424 combinations of speed, object, distance, and geometric parameters. A computer program was used to calculate the hazard index for each of the combinations, to sort the hazard indexes in order of decreasing numerical value, and then provide an ordered listing of the parameters that gave rise to these indexes. Since the combinations were generated theoretically, some of the conditions that appear high on the ordered listing may not exist anywhere along the roadway system. An examination of the top 150 hazard indexes (1.8 percent of the total combinations) for wide rural roads with an ADT of 8000 identified the following characteristics:

1. Forty-five percent of the entries have speeds of 88 km/h and five entries have speeds of 56 km/h.
2. Each type of object appears in the list of the top 150 hazard indexes.
3. Seventy-eight percent of the entries are for grades of less than -5 percent; 80 percent involve curves in excess of 6° curvature.
4. Location on the outside of curves is dominant al-

though locations on the inside of curves and on tangent sections also appear.

5. No objects more than 3 m from the edge of the roadway appear in the list of the top 150 indexes, and most are within 1.5 m.

REFERENCES

1. O. K. Dart and L. Mann. Geometry-Accident Study. Louisiana State Univ., 1968.
2. J. C. Glennon. Roadside Safety Improvement Programs on Freeways. NCHRP, Rept. 148, 1974.
3. J. C. Glennon and C. J. Wilton. Roadside Encroachment Parameters for Nonfreeway Facilities. TRB, Transportation Research Record 601, 1976, pp. 51-52.
4. J. W. Hall. Identification and Programming of Roadside Hazard Improvements. Federal Highway Administration, U.S. Department of Transportation, Interim Rept. FHWA-MD-R-76-2, Jan. 1976; Final Rept., June 1977.
5. J. W. Hall and others. Roadside Hazards on Nonfreeway Facilities. TRB, Transportation Research Record 601, 1976, pp. 56-58.
6. J. W. Hutchinson and T. W. Kennedy. Safety Considerations in Median Design. HRB, Highway Research Record 162, 1967, pp. 1-29.
7. T. E. Mulinazzi and others. Operational Effects of Secondary Stopping and Recovery Areas. TRB, Circular 182, Dec. 1976.
8. G. D. Weaver, D. L. Woods, and E. R. Post. Cost-Effectiveness Analysis of Roadside Safety Improvements. TRB, Transportation Research Record 543, 1975, pp. 1-15.
9. P. H. Wright and L. S. Robertson. Priorities for Roadside Hazard Modification. Insurance Institute for Highway Safety, March 1976.

Publication of this paper sponsored by Committee on Operational Effects of Geometrics.

Macroscopic Modeling of Two-Lane Rural Roadside Accidents

Donald E. Cleveland and Ryuichi Kitamura, Department of Civil Engineering, University of Michigan

A macroscopic study of off-road accident, road, and traffic flow characteristics on the rural two-lane state trunkline system was made to assist the Michigan Department of State Highways and Transportation (MDSHT) in developing priority programs for roadside hazard improvement. Statewide accident data for the period between 1971 and 1974 were analyzed, and, based on these data, a macroscopic modeling effort was undertaken for two-hundred and seventy 3.2-km (2-mile) sections of homogeneous two-lane road that had widely varying road and traffic conditions. Road data came primarily from analysis of MDSHT photolog files. Multiplicative models for different groups of average daily traffic were developed in which restriction on passing-sight distance, number and length of curves, and length of road with exposure to roadside obstacles within given distances from the road were found to be the main explanatory variables. These models, which were evolved dynamically with the aid of statistical computer programs, were tested for the validity of underlying assumptions and were shown to explain as much of the variance as would be expected assuming a Poisson process of accident frequency. The models were validated by using additional data for two cases of low average daily traffic, and satisfactory results were obtained. Several immediate uses for the models are presented.

Despite heavy urbanization, more than one-third of the total automotive accidents reported in Michigan happen on rural roads outside of incorporated areas (1). These accidents occur on facilities that range from low-flow, unimproved routes to multilane, intercity freeways. Even an agency such as the Michigan Department of State Highways and Transportation (MDSHT), which is responsible for the most important 12 900 km (8000 miles) of highway in the state—the portion that carries 38 percent of the rural traffic—has a range in rural facilities from 4.3-m (14-ft) wide two-lane routes to six-lane divided freeways.

This system suffers approximately 50 000 accidents and a total of 600 deaths/year (1). In recent years, much attention has been focused on these accidents in which damage or occupant injury results from the vehicle leaving the road by striking an obstacle or losing its stability and turning over.

Highway agencies have several countermeasures available that can reduce the toll from off-road accidents. Obstacles can be removed or moved farther from the road; they can be weakened so as to break away without damaging the vehicle extensively; and they can be protected by devices that absorb the energy of the vehicle or redirect it along a safer path. In addition, the ground form created by such features as ditches and slopes can be made more forgiving by reshaping and stabilizing it for improved vehicle stability under emergency conditions.

It is recognized that a program of creating a "forgiving road" on every kilometer of the Michigan rural highway system would require a tremendous investment in funds and time. Agencies with rural responsibilities must invest their limited funds and manpower resources in those roadside improvements that return safety benefits that justify these expenditures, and these investments must be made in a sequence that will maximize the time-scaled return to society.

Clearly, a key step in a roadside safety program is to be able to predict what will happen when a roadside improvement of a particular type is made. An organized way of developing the necessary understanding to make such a prediction is to create a model that is accurate enough to be used in the investment decision. Useful models must be able to predict the consequences of a wide range of improvement alternatives. Unfortunately, current understanding of the causes of accidents is inadequate, and only in recent years have sustained modeling efforts been started and promising results obtained (2).

As a part of a sponsored research project for MDSHT, the investigators have explored and reported on the availability of models that are useful in predicting the frequency and severity of off-road accidents on short sections of road over fixed periods of time by using as inputs only knowledge of road and traffic conditions (3). Earlier efforts are found in the literature (4).

A brief summary of the findings of earlier investigations is that off-road accidents are particularly susceptible to occurrence on curves, at locations with restricted sight distance, on gradients, in the presence of structures, and where roadways and traffic flow vary.

The most extensive attempt to model this phenomenon was presented by Foody and Long (5). Their best regression models for predicting single-vehicle off-road accidents involved as many as 14 road and traffic variables and explained only 37 percent of the variance in the accident rate. They found that traffic flow, sight-distance restriction, road geometry transitions, and shoulder width were the most important of these variables. In an additional analysis, they concluded that shoulder width and surface stability were of primary importance in off-road accident experience. They concluded that the relative possible improvement resulting from removal of roadside obstacles was quite small, that the development of such a program would not yield adequate returns, and that attention should be focused on shoulders and the road surface itself. After careful study of the analysis of Foody and Long, it is believed that many possible contributing roadway and traffic elements were not taken into account simultaneously; the obvious existence of interactions among these elements casts serious doubts on the validity of the findings.

GLENNON MODEL

Glennon recently developed a detailed and widely known model that predicts the number and severity of accidents associated with each specific off-road obstacle (2). If the model were completely satisfactory (it is still being refined) and if a highway agency had full information on all roadside obstacles, preferably in an easily retrievable form, the Glennon model could be applied virtually automatically to the entire roadway system, sections that have particular problems could be identified, possibilities for improvement could be determined, and cost-effectiveness analyses could be made. We are not aware of any highway agencies that have data sources in this form and, accordingly, the work presented in this paper is intended to serve primarily as a filtering device by which those highway sections and types that have the greatest potential for off-road accidents can be identified. Then data for the application of the Glennon model can be developed and cost-effectiveness analyses of potential improvements made at the necessary level of detail.

It must be noted that the Glennon model in its most recently available published form does not specifically capture the observed higher frequency of off-road accidents on curves in comparison with tangents; does not respond to other alignment, intersection, or cross-section elements; and maintains that the frequency of accident occurrence is directly proportional to traffic flow, a finding not generally supported by authoritative empirical studies.

METHODOLOGY

The approach used in this research involved two stages of data acquisition. In the first stage, statewide accident data for all two-lane rural roads for the years 1971 to 1974 were obtained from MDSHT. From the accident summaries themselves, some information on the roadway was obtained (curvature, presence of an intersection, type of object struck). Average daily traffic (ADT) was acquired from another state data file. These data were classified appropriately, and statewide effects were determined.

The second stage involved using the same accident files for locating accidents and obtaining information on the roadway, roadside, and traffic from other sources such as studies of sufficiency rating, ADT files, and a detailed engineering study that used the MDSHT photolog system [a photographic record of the driver's view available at each 16.1 m (0.01 mile) along the main trunkline system] to study roadway sections of concern. The main modeling effort was guided by the analysis of first-stage data and used the second-stage data as inputs. A stratified sampling technique was used in determining a set of uniform 3.2-km (2.0-mile) roadway sections.

The modeling effort involved the careful selection of causal variables and alternatives of model structure. The interactive development and improvement of the models, including the estimation of parameters, were undertaken by using the University of Michigan OSIRIS and MIDAS systems (3). Models were subjected to standard tests that followed currently accepted techniques.

After the completion of the modeling effort, it was possible to validate two of the models by using easily available roadway data not used in the processes of modeling or parameter estimation (6). In evaluating the model's predictive performance, a Poisson assumption was postulated as an underlying structure of the accident count on homogeneous sections. At the same time, this assumption was applied to filter out the "outliers" that had extreme accident experience.

RESULTS

Statewide

In the statewide data analysis, it was found that 75 percent of the off-road accidents on rural two-lane highways are of the fixed-object type and the remainder are turnover accidents. Approximately one-third of the fixed-object and three-fifths of the turnover accidents involve injuries or fatalities.

It was found that the off-road accident rate decreases with increasing ADT. Roadway alignment was found to have a dominant effect on the severity of these accidents, and there was a high rate of injury accidents on curves. Furthermore, in the comparison of fixed objects and turnover accidents, there was a higher occurrence of turnover accidents on curves.

It was found that the type of object struck is closely related to accident severity. However, this effect also interacts with roadway alignment in that the severity index (the fraction of all accidents that involve injuries or fatalities or both) is higher on curves for every type of object; this object-alignment interaction with severity is most noticeable for rigid objects with higher indexes of severity. It was also found that the severity of fixed-object accidents is less in intersection areas.

Some of the more significant results of the statewide analysis are shown in Figures 1 through 3 and are given in the table below. Figure 1 shows accident rates versus ADT. It can be seen that accident rates decrease as ADT increases, particularly for turnover accidents. The very high accident rate of the less-

Figure 1. Off-road accident rate and ADT: 1973.

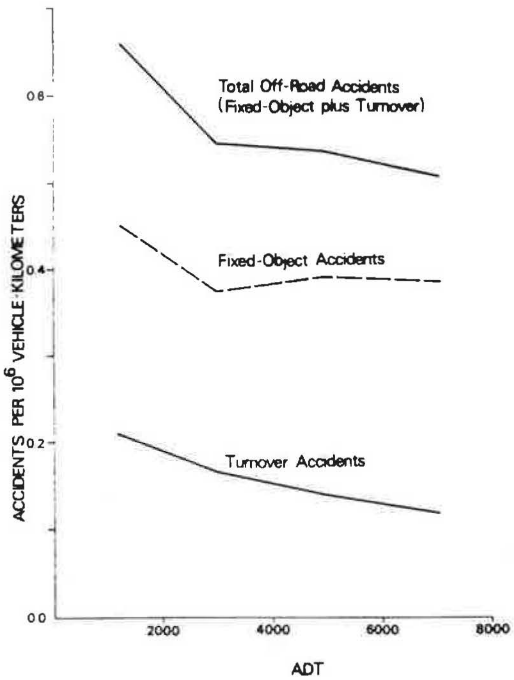
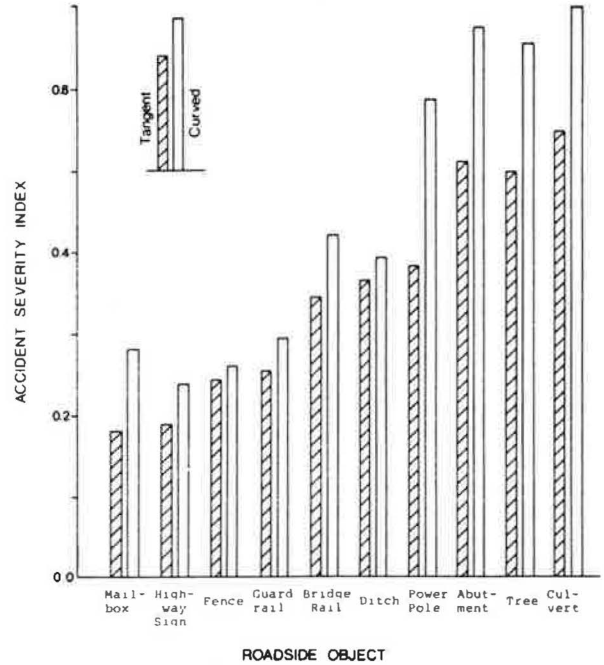


Figure 2. Accident severity indexes of objects struck by alignment.



than-2000-ADT class is particularly noticeable.

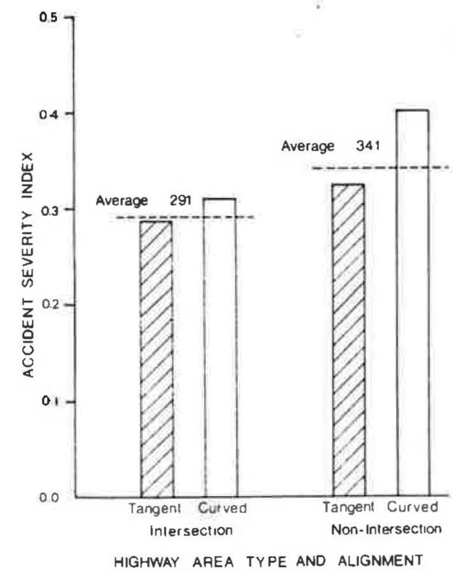
The table below compares the 1974 frequencies of fixed-object accidents along two-lane rural MDSHT trunklines with fixed-object accidents on all rural roadways in the state (which are generally lower type facilities):

Object	Strikings on MDSHT Trunkline Number	Strikings on MDSHT Trunkline Percent	Number of Strikings on All Rural Roadways	Trunkline Percentage of All Rural Roadways
Guardrail	575	15.2	3 148	18.3
Highway sign	448	11.9	2 622	17.1
Power pole	280	7.4	2 806	10.0
Culvert	82	2.2	423	19.4
Ditch	965	25.6	7 803	12.4
Bridge abutment or pier	27	0.7	300	9.0
Bridge railing	43	1.1	382	11.3
Tree	556	14.7	6 085	9.1
Highway or rail-road signal	15	0.4	102	14.7
Building	32	0.9	360	8.9
Mailbox	402	10.7	2 737	14.7
Fence	128	3.4	1 544	8.3
Island or curb	17	0.4	195	8.7
Concrete barrier	12	0.3	328	3.7
On-road object	90	2.4	1 250	7.2
Other off-road ob-ject	80	2.1	689	11.6
Overhead object	19	0.5	90	21.1
Unknown	3	0.1	149	2.0
Total	3774	100.0	31 013	12.2

Objects such as power poles and trees have a lower frequency of being struck along trunklines, which indicates the better clearance of these roadsides. The higher frequencies of striking of guardrails, highway signs, and traffic signals along the trunkline indicate the greater density of these objects along these routes.

The overall average severity index is 0.328. Figure 2 shows the effect of alignment on the index of accident severity as well as variation in the severity index for different types of objects struck. For all

Figure 3. Accident severity index and alignment by intersection.



objects, accident severity on curves is greater than on tangents, and unyielding objects such as power poles, trees, bridge abutments, and culverts have much higher severity indexes on curves. The differential effect of roadway alignment on severity is compared in Figure 3 by proximity of an intersection. The non-intersection areas generally have accidents of greater severity than the intersection areas. Furthermore, the effect of alignment on severity is not as great in intersection areas. Clearly, the value of the severity index used in object hazard evaluation must respond to roadway alignment, especially for objects that show a high-severity difference.

Accident Prediction Modeling

Since the input data for an operational model would be developed for a highway agency by using data from its own files, the models described in this section are based on MDSHT accident and highway files.

There is a tremendous variation in the frequency of off-road accidents on different sections of the Michigan trunkline system. Figure 4 shows this variation for the sample of two-hundred and seventy 3.2-km (2-mile) sections for the 4-year period between 1971 and 1974, which was used as a basis for the modeling effort. It can be seen that 10 sections recorded more than 20 off-road accidents in this period, one section had 34, and 26 sections had no reported off-road accidents.

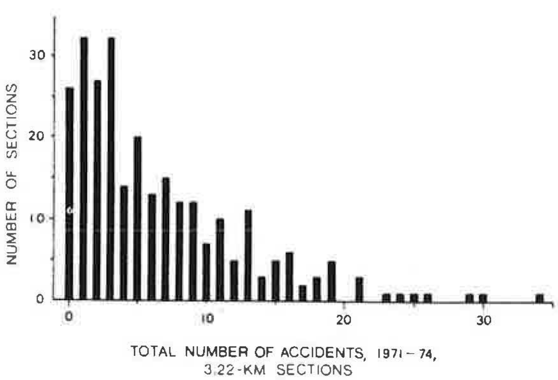
Models were developed separately for total off-road accident experience and for injury-fatality accident experience. Fixed-object accidents were modeled separately from turnover accidents because of their different characteristics. Because of an anticipated possible effect of the national 80-km/h (55-mph) speed limit (effective in March 1974 in Michigan), data for the period from 1971 to 1973 were initially modeled separately from the 1974 data. The results, however, showed that there was no important difference, and data for the entire 4-year period were then pooled and used in all subsequent modeling efforts.

It should be noted that this research concentrated on the occurrence of accidents and not on the accident rate. It is believed that the ultimate figure of merit is the number of accidents and that the use of rates can mask this effect. Since the models in this research are of a macroscopic nature, a decision was made to deal with a fixed length of highway (3), and only variables that summarize the relevant highway and traffic characteristics of such a section were used as inputs to the model.

The first task in the modeling effort involved the identification of relatively easily obtainable data on variables that were expected to be causal or strongly associated with the occurrence of off-road accidents. The table below gives the variables that were used in the analysis (1 m = 3.3 ft):

Variable	Abbreviation Used
Area	AREA
Pavement width	PAVE. W.
Shoulder width	SHOULD. W.
Percentage sight restriction	PSR
Rolling	*
Number of curves	NC
Percentage of segment length curved	PCL
ADT	ADT
Number of intersections on curves	NIC
Number of intersections on tangent	NIT
Total number of intersections	NITO
Shoulder treatment	*
Ditch condition	DITCH
Object stiffness	STIFF
Percentage exposure length to objects within 2 m	OB6
Percentage exposure length to objects within 3 m	OB10
Percentage exposure length to objects within 4 m	OB14
Percentage exposure length to objects within 6 m	OB20
Percentage exposure length to objects within 9 m	OB30

Figure 4. Distribution of total number of accidents on 270 roadway sections: 1971 to 1974.



Variables represented by an asterisk did not appear in the results. Traffic flow was represented by the 1971-1973 average of three ADT values (1974 data were not available).

Variables expected to be associated with alignment included the percentage of the section with passing-sight restrictions, a characterization of the terrain as rolling or level, a count of the number of curves in the section further broken down by the presence or absence of intersections on curves, and the total length of curved road in a section. Measures associated with the cross section and roadside included the width of pavement and shoulder, the type of stability provided by the shoulder treatment, the predominant distance to drainage ditches and a description of the cross-sectional abruptness of these ditches, the exposure distance to obstacles within a variety of distances of the edge of the roadway, and a characterization of the energy exchange characteristics of those obstacles located less than 4.2 m (14 ft) from the edge of the roadway. The photolog study provided much of the above information.

The next step involved drawing a probability sample of rural 3.2-km (2-mile) sections for study. The initial task was to identify the population of two-lane rural MDSHT trunklines in the state. An initial screening was made of the 1974 MDSHT sufficiency rating report. At later stages of the process, additional sections were eliminated, primarily because of the discovery of sections in urbanized villages classified as rural, sections that had been reconstructed to multilane standards, and those at the approaches to urbanized areas. A total of 1392 rural two-lane segments were identified. The strata formed for the final sampling consisted of three areal classifications for the state (the Upper Penninsula is much more rugged, rural, forested, and less densely populated than the highly urbanized southern sections), four ADT classification groups, four classifications of shoulder width, three classifications of pavement width, and the percentage restriction on passing sight distance and the general terrain classification of the section. If sections with all combinations of each stratum existed, there would be about 1400 possibilities.

Next, a review of individual sections was made to ensure that the length was 3.2 km (2 miles) or greater. Some shorter sections, frequently those with high ADTs near urbanized areas, were eliminated from the sample population. For each section, a random point of beginning was selected, and the succeeding 3.2 km were used.

It was then determined that the availability of time and funds limited the main data-acquisition effort (photolog analysis) to between 250 and 300 sections. This meant that an approximately 20 percent sampling rate of all sections could be used, which resulted in a slightly less than 10 percent sample of total rural MDSHT two-lane highway.

It was decided that stratified random sampling would be used since it is of crucial importance to obtain information on all existing combinations of possibly contributing causal elements. All strata that had two or fewer sections were selected for the final sample. For the other strata at least two sections were included in the sample. Combinations that involved extreme values of the strata were overrepresented. This sample particularly protects the results from extrapolation errors in the use of the resulting model at the possible sacrifice of accuracy in the most frequently occurring combinations.

At the conclusion of the sampling, a total of 270 sections had been identified and studied. Thus, the modeling efforts for this study are based on data from this 869.4 km (540 miles) of Michigan trunklines.

The next step was the use of the automatic interaction detection (AID) multivariate analysis technique, an extremely useful screening method developed at the University of Michigan (7). An effective method of presenting the results of an AID analysis is a branch diagram from which one can see the way explanatory variables interact as well as the importance of individual variables in the explanation of variation, an important early step in the construction of models. One of the AID diagrams used in the research is shown in Figure 5.

Although the average number of turnover accidents between 1971 and 1974 on the 270 sections was 1.91, it is obvious that traffic flow (ADT), the fraction of the road that is curved (PCL), the length of the route that has fixed objects relatively close to the pavement (OB14 and OB20), and the fraction of the road that has inadequate passing sight distance (PSR) affect this average immensely. Although sections that have an ADT less than 500 average only 0.28 accidents, those that have high ADTs, much curvature, and fixed objects within 6 m (20 ft) of the edge of the pavement along much of the route average 6.23 accidents. It should also be noted that this simple, unstructured model explains more than 42 percent of the variation in the entire data set.

The next step was to develop an appropriate model by using multiple regression techniques and the AID results. The AID process signaled the necessity of stratifying the models when clearly different variables were involved. The regression model structures explored included both linear and multiplicative forms. However, because the analysis indicated the superiority of the multiplicative models over the linear models, the linear models are not described here (3).

Total Accident Estimation Models

The AID branch diagram for total off-road accident experience is shown in Figure 6. This stratification accounts for 76 percent of the variation in only 18 ultimate classes of predictive variable combinations. The average number of accidents ranges from 1.08 for roads with low ADT and good passing sight distance to 26.0 for curved sections with many fixed objects within 6 m (20 ft) of the surface and ADTs greater than 7000 vehicles/d. The stratification is dominated by ADT, and a review of the variables for each of the ADT groups led to a decision to model separately each of the four ADT groups shown in Figure 6.

The final estimating equations for the four ADT classes are given below. The dependent variable y is always the number of accidents in a 3.2-km (2.0-mile) roadway segment for a 4-year period. PSR, PCL, and OB20 take values that range from 0 to 100.

For ADT < 750,

$$y = 0.024(ADT)^{0.70}(PSR + 1)^{0.18} - 1 \quad (1)$$

where t-statistics for the coefficients are 2.84, 2.98, and 3.03 respectively; $R^2 = 0.34$; and N = 50. For 750 ≤ ADT < 1500,

$$y = 2.54(PSR + 1)^{0.24} - 1 \quad (2)$$

where t-statistics for the coefficients are 5.41 and 4.38 respectively, $R^2 = 0.26$, and N = 58. For 1500 ≤ ADT < 3500,

$$y = 0.016(ADT)^{0.69}(PCL + 1)^{0.068}(OB20 + 1)^{0.29} - 1 \quad (3)$$

where t-statistics for the coefficients are 2.11, 2.68, 1.97, and 5.16 respectively; $R^2 = 0.32$; and N = 82. For ADT ≥ 3500,

$$y = 0.12(ADT)^{0.46}(NC + 1)^{0.35}(OB20 + 1)^{0.21} - 1 \quad (4)$$

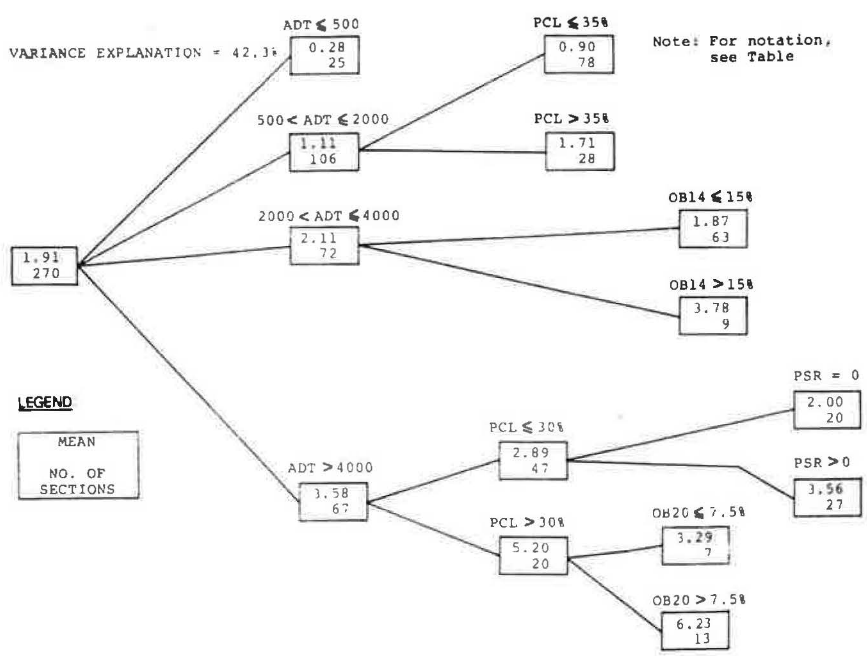


Figure 5. AID branch diagram: 1971 to 1974 turnover accidents.

Figure 6. AID branch diagram: 1971 to 1974 total accidents.

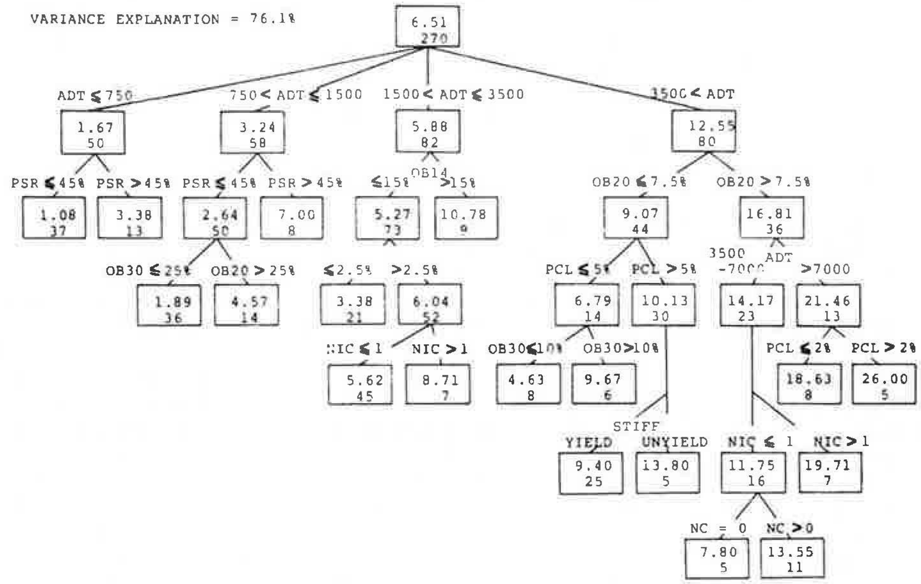
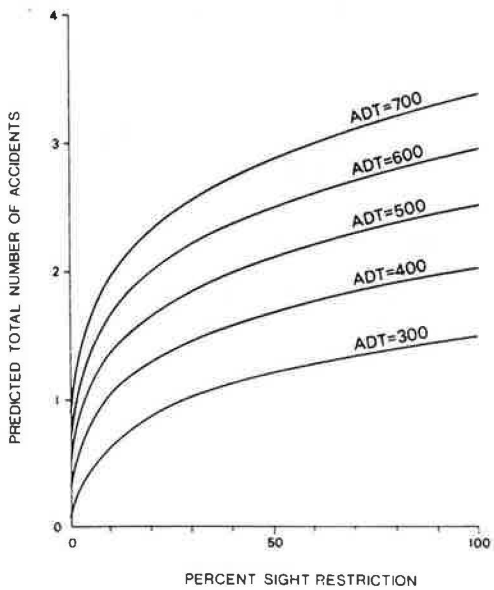


Figure 7. Model for prediction of total accidents: < 750 ADT.



where t-statistics for the coefficients are 1.91, 3.38, 4.81, and 4.66 respectively: $R^2 = 0.49$; and N = 80.

The variance explanation by these models may appear low. However, it should be noted that about 70 percent of the variance has already been explained by the ADT stratification. The entire variance explanation by these models exceeds 82 percent. It is seen that ADT, restriction on passing sight distance (PSR), length of route that has obstacles within 6 m (20 ft) (OB20), percentage of the road that is curved (PCL), and number of curves (NC) are the variables that appear in these models. As an example of the simplicity of the relations, Figure 7 shows a plot for sections that have an ADT of 750 vehicles or less. The lessening effect of increasing ADT, even at this low level, is clear, as is the importance of good alignment.

Estimation Model for Injury and Fatal Accidents

The AID branch diagram for injury and fatal accidents is shown in Figure 8. Note that the effect of ADT is not so dominant for injury accidents as for total accidents. Although curved alignment is important in the prediction of injury accidents, the variable for restriction on passing sight distance does not appear at all. It appears that the injury accident is more sensitive to horizontal alignment than the less severe accident and that vertical alignment, an important component of the passing sight restriction, is less important in injury accidents. This result is consistent with the statewide results described earlier.

One interaction that involves pavement width should be noted. The diagram shows that on high-ADT roadways that include much exposure to roadside objects and lengthy curved sections, the 6.0-m (20-ft) wide and 6.6-m (22-ft) wide surfaces have 1.5 times as many injury accidents as do 7.2-m (24-ft) wide pavements.

The injury-fatality model is given below. The model predicts y, the number of injury-fatality accidents in a 3.2-km (2-mile) roadway segment for a 4-year period:

$$y = 0.039(\mathrm{ADT})^{0.52}(\mathrm{PCL} + 1)^{0.096}(\mathrm{OB10} + 1)^{0.069}(\mathrm{STIFF}) - 1 \qquad (5)$$

where t-statistics for the coefficients are 11.94, 14.98, 4.70, 2.36, and 2.12 respectively; $R^2 = 0.49$; N = 270; and STIFF assumes a value of 1.36 if unyielding objects exist within 4.2 m (14 ft) of the edge of the pavement and 1.17 otherwise.

It is seen that injury-fatality accident prediction is approximately proportional to the square root of the ADT, higher roots of the fraction of the road that is curved, and the length of road that has objects closer than 3 m (10 ft). There is an approximately 17 percent effect for the energy exchange characteristics of the obstacles within 4.2 m of the road. The presence of objects within 3 m in this model suggests that the number of injury accidents is more affected by closer objects. This model can be viewed as a macroscopic version of the Glennon model in which a term is added to capture the effect of alignment.

Summary of Variables

A count of the frequency of explanatory variables was made based on the four primary AID analyses. It was found that ADT was always the most important variable and appeared more than twice as frequently as any other

Figure 8. AID branch diagram: 1971 to 1974 injury-fatality accidents.

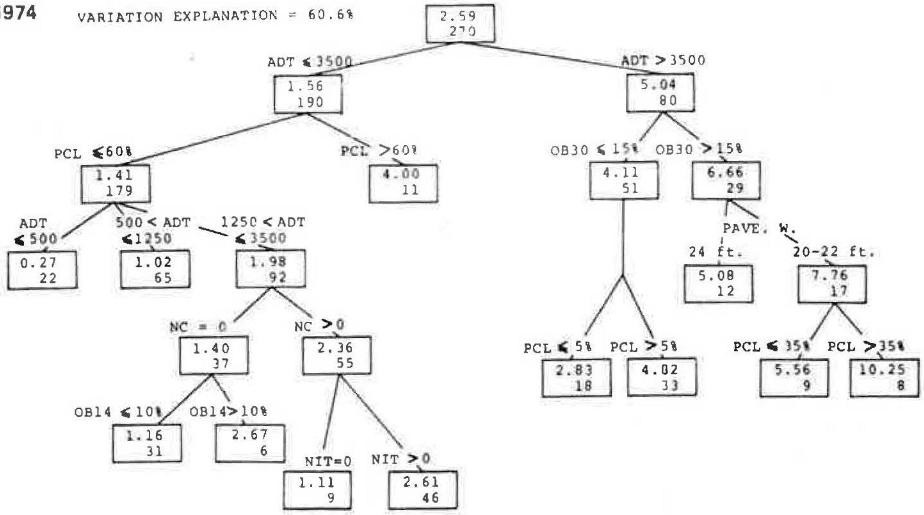


Figure 9. Validation check for sections that have <750 ADT.

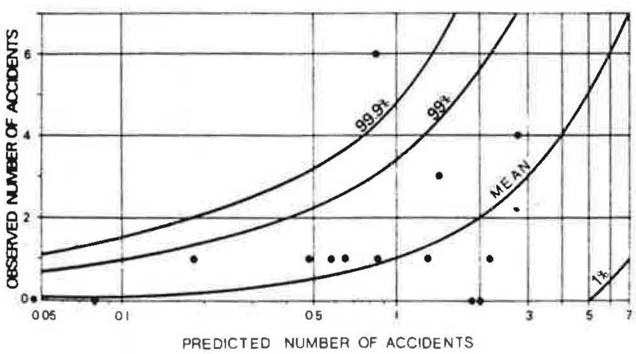
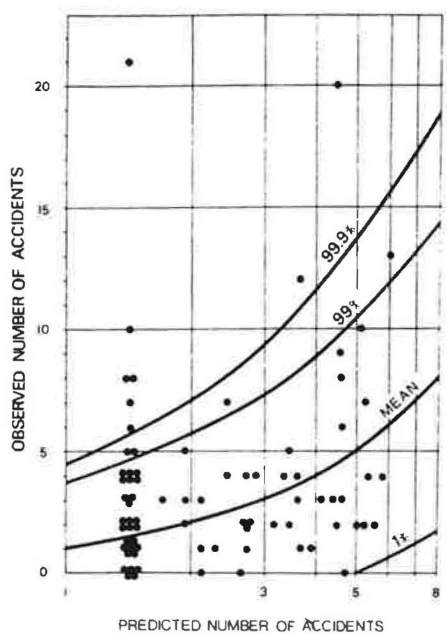


Figure 10. Validation check for sections that have ADT from 750 to 1500.



explanatory variable. Other variables that appeared frequently were the length of road that had obstacles within 4.2 m (14 ft), 6.0 m (20 ft), and 9.0 m (30 ft) of the pavement edge and the fraction of the route that was curved and had inadequate passing sight distance. All other measures were of lesser importance, and variables that represented obstacles very close to the roadway [less than 2 m (6 ft)], shoulder treatment, rolling terrain, and number of intersections did not appear in any of the AID results.

VALIDATION OF THE MODEL

A partial validation study of the effectiveness of the models was developed. An extensive validation was not feasible within the time and fund constraints of the project. However, it was possible to use readily available data for ADT and restriction on passing sight distance to determine the effectiveness of the low-ADT total accident prediction models. The data required to compare 14 sections that had ADTs of less than 750 vehicles and 78 sections with ADTs of from 750 to 1500 vehicles that were not used in the model formulation and calibration were developed (5). The results of this analysis are shown in Figures 9 and 10 where actual accident experience is shown versus the predicted number of accidents.

Three Poisson probability bounds are drawn in these figures. These bounds imply that, if a highway section belongs to the population that the accident estimation model represents and if the Poisson law describes the distribution of the number of accidents, the observed number of accidents in the section should fall within these bounds with the probability associated with the bounds. Then, for those sections that lie outside the bounds, one may conclude that they have accident expectations that are different from those indicated by the model. It is reasonable to expect that some factors other than those that significantly affect the accident experiences of most of the sections included in the model are involved with these outliers. The generally good fit for most of the sections can easily be seen. However, a number of locations (10 to 20) are clearly out of control.

DISCUSSION AND APPLICATION

In the models developed in the research, the exponent of ADT is always less than 1, which confirms the diminishing effect of ADT on accident occurrence found by other observers. Clearly, obstacle-hazard evaluation models, such as Glennon's, should take this effect into account. Furthermore, the accident prediction models have supported and further quantified the importance of roadway alignment on off-road accident occurrences. This is another area in which Glennon's original model requires further development. As the statewide analysis indicated, the effects of alignment are twofold: those on accident occurrence and those on severity.

The importance of the roadway cross section was not supported by the models as it was in the Ohio results. Accordingly, we cannot support a belief in the importance of a shoulder stabilization program for Michigan highways as a means of counteracting the off-road accident or its severity.

These prediction models can be used as filtering devices in defining highway sections that have high accident rates and where more detailed microscopic studies should be made. The advantage of this filtering approach is clear from Figure 10. A simple ordering of sections according to accident frequency does not necessarily provide a set of sections that have higher accident rates than normally expected. Note that many sections that have high accident rates are within reasonable Poisson bounds. Particular attention should be given to an engineering analysis of the sections outside the 99.9 percent bound region as well as to all sections whose fundamental characteristics predict a high rate of off-road accidents. Another use of the models is the preliminary evaluation of programs for the removal of roadside objects or overall evaluation of systemwide accident improvement potentials.

CONCLUSIONS

This research has shown that a small number of carefully selected, straightforward causal variables can be combined in a multiplicative mathematical model to explain as much of the variability in rural, two-lane, off-road accident frequency as could be expected. The models are usable directly to identify locations that have the highest probability of frequent off-road accidents as well as to point out those locations where additional factors may be at work and engineering study is clearly needed.

ACKNOWLEDGMENTS

The cooperation of Donald E. Orne, Maurice Witteveen, William Lebell, and Preston Masters of the Michigan Department of State Highways and Transportation in the sponsored portion of this work is gratefully acknowledged. James Yao-Ching Chang of the National Cheng Kung University, Tainan, Taiwan, assisted in the validation phase of this effort.

Much of this research was conducted under contract with the Michigan Department of State Highways and Transportation, and the opinions, findings, and conclusions expressed in this publication are ours and not necessarily those of the Michigan Department of State Highways and Transportation or the University of Michigan.

REFERENCES

1. Michigan Traffic Accident Facts. Michigan Department of State Police, Lansing, Dec. 1976, 16 pp.
2. J. C. Glennon. Roadside Safety Improvement Programs on Freeways: A Cost Effectiveness Priority Approach. NCHRP, Rept. 148, 1974, 64 pp.
3. D. E. Cleveland and R. Kitamura. A Study on Two-Lane Rural Roadside Accidents. Department of Civil Engineering, Univ. of Michigan, Ann Arbor, Aug. 1976, 183 pp.
4. J. K. Kihlberg and K. J. Tharp. Accident Rates as Related to Design Elements of Rural Highways. NCHRP, Rept. 47, 1968, 173 pp.
5. T. J. Foody and M. D. Long. The Identification of Relationships Between Safety and Roadway Obstructions. Columbus Bureau of Traffic, Ohio Department of Transportation, Jan. 1974, 83 pp.
6. J. Y. Chang, D. E. Cleveland, and R. Kitamura. Validation of a Two-Lane Trunkline Roadside Accident Prediction Model. Department of Civil Engineering, Univ. of Michigan, Ann Arbor, 1977, 6 pp.
7. J. A. Sonquist and others. Searching for Structure. Institute for Social Research, Univ. of Michigan, Ann Arbor, 1971.

Discussion

John C. Glennon, Transportation Consulting Engineer, Overland Park, Kansas

I would like to commend the authors for their work. Their report contributes to the state of the art in the area of roadside safety and raises other questions that need to be answered. Perhaps the most significant finding is that the severity of fixed-object accidents is higher on highway curves than on highway tangents for all objects. Given that encroachment rates are also higher on curves, this would suggest that the rate of off-road injury and fatal accidents on curves is an order of magnitude higher than the rate on tangents.

I am surprised that the authors either did not review or at least did not reference the Federal Highway Administration (FHWA) report in which I modified my model to account for roadside hazard for two-lane roads (8). Reference to that report indicates that, contrary to the authors' statement, the most recently available inputs to the subject model with regard to two-lane highways do account for a decreasing off-road accident rate with increasing ADT. The other item of interest in comparing the two researches is that the severity indexes found in the FHWA research tend to substantiate those found by the authors.

The second part of the paper attempts to develop methods (models) for identifying priority highway sections for roadside safety improvements. Although the authors made a commendable effort, they seem to have performed one more in a long line of unsuccessful multivariate analyses aimed at relating accident occurrence to roadway and traffic variables. The only variable that explained a substantial portion of the accident variance was traffic volume. But this conclusion is not a new one.

Although the modeling results may provide some general guidance in judging the relative roadside hazard of highway sections, the statistical practicality of these results must be viewed with some skepticism. For example, consider the validation plots shown in Figures 9 and 10. In Figure 9, the model for 750 ADT or less only predicts accident occurrence within ±50 percent for about one-third of the validation sites. Figure 10 has a slightly better result, but this model (for ADTs of from 750 to 1500) still only predicts accident occurrence within ±50 percent for about 43 percent of the validation sites. In addition, for many of the outliers the prediction equation is more than 100 percent in error. These results are not encouraging in terms of the reliability of predictions.

Perhaps the lack of model precision lies in the abstract nature of the selected variables. The percentage of passing sight restriction is a good example. Passing sight distance as defined can only be related to traffic operations in a general sense as demonstrated by the widely different treatments found in the 1965 blue book of the American Association of State Highway Officials (9) and the Manual on Uniform Traffic Control Devices (10). But the percentage of roadway that has restricted passing sight distance is one level of abstraction farther removed. For example, what is the effect on traffic operations of areas of acceptable passing sight distance that are not within legal passing zones? In a similar sense, using the percentage of roadway on curves without regard to the specific geometrics of those curves and their longitudinal relationship to each other presupposes an abstract effect on off-road accidents that may, in fact, be nonexistent.

In conclusion, I again commend the authors for their research on a very difficult problem. Their work has provoked some new thoughts for me and I hope for others as well.

REFERENCES

8. J. C. Glennon and C. J. Wilton. Effectiveness of Roadside Safety Improvements. Federal Highway Administration, U.S. Department of Transportation, Research Rept. FHWA-RD-75-23, Nov. 1975.
9. A Policy on Geometric Design of Rural Highways—1965. AASHO, Washington, DC, 1966.
10. Manual on Uniform Traffic Control Devices for Streets and Highways. Federal Highway Administration, U.S. Department of Transportation, 1971.

Thomas E. Mulinazzi, University of Maryland

The need for models such as the ones presented by Cleveland and Kitamura is great. Highway agencies are experiencing financial austerities, so that each dollar spent must be justified. The 1973 Federal-Aid Highway Act established "new categories of earmarked funds for three roadway-related safety programs on Federal-aid highways other than the Interstate System: protection of railroad-highway grade crossings, improvements at high-hazard accident locations, and elimination of roadside obstacles." However, the 1976 Federal-Aid Highway Act combined the high-hazard location and roadside obstacle programs. The end result of combining these two programs has been the virtual elimination of the roadside obstacle program on a systemwide basis. A project under the roadside obstacle program cannot compete on a benefit-cost basis with the typical project under the high-hazard location program. Therefore, the highway agencies must decide on which high-hazard locations (including some rural roadside accidents) to include in a safety improvement program.

This is not as easy as it seems. The technology exists to solve many of these problems, but the data base on which to develop a priority system is weak at best, and the money to implement the improvements is scarce and in tough competition with other highway projects. The current roadside safety improvement program is more of a reactive (after the accidents occur) than an active approach. I am strongly in favor of the preventive maintenance approach or the active approach to reducing highway accidents and accident severity. This is where models such as those presented by Cleveland and Kitamura have an application. They can be used to indicate the accident potential of a section of rural highway before the accidents occur.

The models developed in the paper were based on reported off-road accidents. It is my opinion that a majority of the vehicles that leave the roadway are not involved in reported accidents. This could indicate, however, that these unreported departures occurred on forgiving roadsides and that the reported departures took place on highway sections that need roadside safety improvements. So the unreported accident situation may not be a significant problem in determining the accident potential of a highway section.

I like the fact that 3.2-km (2-mile) sections were used to develop the models. Using road sections instead of specific locations reduces the effects of improper reporting of accident location. The accident data base is the weak link in developing any model for highway accident potential.

In a 1977 report of the Federal Highway Administration (11), various design elements—such as degree of horizontal curve, type of curve transition, superelevation rate and runoff, sight distance, and grade—were to be evaluated to determine the influence of each on highway accidents. The main problems were the lack of independence between criteria and the lack of consideration for consistency in design elements. This latter point is difficult to include in any model, but it may be significant in determining the accident potential of a road section. For example, a 4° curve in the middle of a winding road may be a safe design element, but a 4° curve at the end of a 14.5-km (10-mile) tangent segment could be, and probably is, a hazardous location even though it is the only sight restriction in the 3.2-km (2-mile) section.

I feel that a strong point of the models developed in this paper is the fact that all the variables but one are very easy to obtain from plans or field inventories. Percentage sight restriction (PSR), percentage of curved length (PCL), number of curves per 1.6 km (1 mile), and an object stiffness factor (STIFF) are readily obtainable. However, the variables, which are based on the percentage length of exposure to objects within a certain distance of the roadway, would be a judgment value in many cases. There is no problem with measuring the length of guardrail, but how would values be determined for this variable if 50 isolated trees were located 6 m (20 ft) from the roadway on a 3.2-km (2-mile) section?

The main problem I have with modeling two-lane rural roadside accidents is the low frequency of accidents on any particular section. I do feel, however, that the approach taken by Cleveland and Kitamura is the first step in developing a roadside safety improvement program. The results of the models will indicate those road sections that could have an accident problem and that require a microscopic engineering analysis. If only highway agencies had the manpower and money to carry out this preventive approach to the roadside safety problems on non-access-controlled rural roads, these models, or models like them, would be worthwhile.

REFERENCE

11. Effectiveness of Design Criteria for Geometric Elements. Federal Highway Administration, U.S. Department of Transportation, RFP-371-7, 1977.

Authors' Closure

We thank both Glennon and Mulinazzi for reading our paper and for their highly relevant and well-chosen comments.

We share Glennon's estimate of the importance of the finding that the severity of fixed-object accidents is much greater on highway curves than on tangents. If nothing followed from this research other than the direction of particular attention to objects located on curves, we would feel that our efforts have been more than worthwhile. It appears that the profession is now at a stage to use the Glennon modified microscopic model for a wide range of applications.

There was some concern with the variables that we used to capture the obviously important roadway alignment effect. In the model-building effort, attention was necessarily paid to variables that could be directly retrieved by our highway agency. In our opinion, the variables we selected meet this criterion for Michigan. Extensive efforts were made to select the best alignment variables from among those available (Table 2). Mulinazzi suggests the need for a quantitative measure that represents longitudinal changes in roadway alignment. Such a measure would also serve as a guideline for consistent roadway geometric design. We agree and would like to have developed such a measure.

The investigators would like to have had much more detailed information on roadway and traffic characteristics available in machine-retrievable form. Unfortunately, the state of practice and economics have not permitted the development of data systems in which obviously better variables are available. On the other hand, it is believed that the variables that we have used provide significant guidance with respect to the type of data file that would be valuable in future data systems.

Concern was also expressed about obtaining data on the length of exposure to objects at various distances from the edge of the road. In the study, these variables were developed by recording the dimensions and offset of the object from the roadway from the photolog screen and then converting them into equivalent exposure length at the edge of the roadway by using Glennon's relation (2). Although this process is time consuming, use of the photolog system eliminates expensive field trips, and developing this measure for the entire roadway system is, for Michigan, not a difficult task.

Concerning the predictive performance of the model, Glennon points out that our models predict the number of accidents on a section within a 50 percent error only one-third of the time. However, attention must be paid to the stochastic nature of accident occurrence, particularly on the low-ADT highways on which the validation studies were conducted. The percentage of predictions within a given percentage of error does not apply as an appropriate criterion to judge model performance. We suggest using an evaluation that involves the total number of accidents predicted on several sections versus those that actually occur and also paying attention to the extreme values. For the <750-ADT group, the total number of observed accidents in the 14 sections used in the validation study was 20 whereas the predicted total was 13.3—a 67 percent error. However, if we eliminate the (to us) obvious outlier, these figures become 14 observed versus 12.6 predicted, clearly a reasonable and unimportant difference. In similar fashion, an even better fit was found for the model for the higher ADT class.

Significant progress has been made during this decade in the identification of locations where off-road accidents are likely to occur as well as in the techniques of counteracting this serious highway safety problem. We are pleased to join our discussants in making some contribution to this effort.

Publication of this paper sponsored by Committee on Operational Effects of Geometrics.

Two-Way Left-Turn Lanes: State-of-the-Art Overview and Implementation Guide

Zoltan Anthony Nemeth, Department of Civil Engineering, Ohio State University

The results of a research project to synthesize existing information on continuous two-way left-turn median lanes and to conduct before-and-after studies to evaluate the effectiveness of such lanes as an access control measure are presented. Recommendations were prepared for the traffic engineer concerned with the evaluation of a situation in which a two-way left-turn median lane is a potential solution to existing capacity and safety problems. The research approach included studies in three distinct areas: a nationwide expert opinion survey, a literature review, and before-and-after field studies. Both the literature review and the survey indicated that two-way left-turn median lanes work well in spite of a wide variety of methods of signing and marking. There is uniform agreement that these lanes have excellent safety records; specifically, head-on collisions in the lanes are extremely rare. The before-and-after studies demonstrated that the effectiveness of the lanes and public reaction depend on proper engineering. A step-by-step decision-making strategy has been developed for the implementation of two-way left-turn median lanes.

To increase efficiency, conserve energy, and reduce air pollution, it is national transportation policy to make maximum use of the available transportation capacity in the existing transportation network. There is a continuing emphasis on transportation system management (TSM) plans designed to solve short-range urban transportation problems. Typical examples of TSM actions are innovative traffic engineering measures that improve both capacity and safety and require a minimal investment of manpower, material, or capital.

The two-way left-turn median lane (TWLTL) falls into this category. It serves to reduce the particular conflict observed on roadways that were originally intended predominantly to serve the through-movement function but are now being called on to satisfy an increasing demand for accessibility as well because of changes in adjacent land use. However, in spite of the increasing use of TWLTLs in recent years, spurred in part by the federal funding provided for such improvements under the TOPICS program and then more recently by other categories of federal-aid funds in urbanized areas, considerable skepticism remains regarding TWLTLs. One major concern is the potential hazard created by permitting two-way movement of traffic in a single median lane.

The objective of this study is to bring about a wider application of the TWLTL by lessening the prevailing uncertainties regarding the effectiveness as well as the proper application of this device. To achieve this objective, existing information on TWLTLs was synthesized, and a questionnaire survey, personal interviews, and a literature review were conducted.

SURVEY OF EXPERT OPINION

Questionnaire

The purpose of the questionnaire survey was to elicit pertinent facts and expert opinions from transportation engineers who had practical experience with TWLTLs. Primary areas of interest were (a) the effect of the TWLTL on traffic safety, (b) the effect of the TWLTL on traffic flow characteristics, and (c) conditions conducive to the installation of such a median lane. Secondary interests were signing and lane-marking practices, optimum lane width, proper use by drivers, police enforcement, public acceptance in general, and cost-effectiveness.

The questionnaire developed contained 16 questions. Of the 90 questionnaires mailed out, 70 were returned; they represented 36 states and one Canadian city. The more significant results are summarized below by subject matter.

Practical Experience

In terms of time, the experience of the 70 respondents was as follows:

Years of Experience	Number of Responses
1 or less	7
1-5	36
5-10	21
More than 10	2
Not reported	4

In terms of the number of TWLTLs, the breakdown of experience was as follows:

Number of TWLTLs	Number of Responses
1	11
2-10	42
10-30	6
More than 30	9
Not reported	2

Effect on Flow Characteristics

Respondents were asked if in their opinion TWLTLs were successful in reducing travel time and friction. The response was as follows:

Increase in Quality	Responses (%)
Significant	77
Little	20
None	3

Ten of the respondents stated that they had conducted studies to support their answers. Five responses were received to the question that asked what factors among the following contributed to the ineffectiveness of less successful TWLTLs: too many left turns, narrow lane, lane markings not maintained, adverse news media, or no enforcement.

Effect on Safety

Respondents were asked if TWLTLs improved the safety of roadways:

Effect on Safety	Responses (%)
Significant improvement	66
Slight improvement	27
No improvement	7
Decrease	0

As might be expected, those who did not perceive any improvement in safety belonged to the one-TWLTL category in terms of experience. Of the 22 respondents who had conducted studies to support their answer, 21 cited significant improvement and one slight improvement.

Public Reaction

The following responses regarding public reaction to TWLTLs were received:

Reaction	Responses (%)
Favorable	62
Mixed	25
Unfavorable	1
No reaction	12

About 10 respondents mentioned that the public found the signs confusing.

Proper Use

Of the 70 responses received, 56 noticed improper use, including 50 who judged the problem severe enough to warrant enforcement. In 30 cases, police enforcement resulted.

Signing, Pavement Markings, and Lane Width

The following responses were received on the use of various types of signs by respondents:

Type of Sign	Responses (%)
Overhead	50
Roadside	34
Pavement arrows	48

Eight percent of the respondents stated that they followed the lane markings recommended in the Manual on Uniform Traffic Control Devices (MUTCD) (18). Four percent used dashed yellow outside, solid yellow inside. Other deviations included white lines, solid yellow lines, or double dashed yellow lines. Lane width ranged from a 2.4-m (8-ft) minimum width to a 5.1-m (17-ft) maximum effective width. The distribution of answers is represented by the mean and mode values given below (1 m = 3.3 ft):

Width	Mean (m)	Mode (m)
Minimum allowable	3.1	3.0
Optimum	3.7	3.6
Maximum allowable	4.4	4.2

Conditions Conducive to TWLTLs

Respondents were asked to name the conditions under which TWLTLs would be most useful. The following factors were mentioned most frequently:

Condition	Number of Responses
High number of driveways per block	42
Commercial development	36
Substantial midblock left turns	18

Perceived Effectiveness Versus Signing

It was hypothesized that there might be a relation between the level of signing of a TWLTL and the effectiveness of the lane perceived by the traffic engineer. Responses received are given in Table 1. The table indicates (if only informally, since no statistical testing was done) that, as level of signing increased from no signs to a combination of all three signs, so did the perceived improvement achieved by the TWLTLs.

Personal Interviews

Selected state, county, and city traffic engineers were interviewed in California, Ohio, Texas, and Washington to discuss design and operational aspects of TWLTLs and to obtain unpublished reports, guidelines, and before-and-after data. Most of the material covered by the literature review was obtained through these personal contacts.

LITERATURE REVIEW

The literature was searched for three types of information: (a) general criteria for application, (b) design details, and (c) evaluation. The objective was to get a general consensus of opinion and an overview of the state of the art.

General Criteria for Application

Seven categories of factors that have been considered in connection with TWLTLs either as warranting factors or as constraints were identified.

Adjacent Land Use

Strip commercial development was identified throughout the literature as the adjacent land use most applicable to use of TWLTLs. Continuous high-density commercial land use of this type is most common in the traffic conditions for which TWLTLs are most effective. But successful applications were also documented in residential areas, commercial-residential areas, and even in industrially developed areas under the proper traffic flow conditions. The applicability of the TWLTL is thus a function of the particular traffic conditions that result from adjacent land use rather than a function of the land use itself. In partially developed areas, the TWLTL will generate more strip development. If this is undesirable, preference should be given to raised medians (2).

Access Conditions and Requirements

Existing access conditions are not easily classified or quantified since there are many ways in which access can be provided and a number of factors that determine the ease of access to fronting properties.

References to existing access as a general warrant are common in the literature (2, 3, 4), and, in most cases, its relevance is ascribed to the extent to which alternative means of access are provided (5, 6, 7). Access gained by negotiating a midblock left turn creates the specific traffic conflict and through-movement delay that TWLTLs are designed to treat; therefore, the availability of access through alternative means, such as parallel streets or alleys, service roads, off-street parking facilities, and U-turn or around-the-block movements, must be considered an important factor in weighing TWLTL proposals versus more restrictive left-turn control measures.

The total access requirement, expressed in terms of midblock left-turn demand, would be expected to present a prime factor for consideration in installing TWLTLs. The literature expresses the importance of this access demand as a general TWLTL warrant but only to the extent that a "high" demand contributes to the "general" traffic conditions that warrant consideration of TWLTLs.

Table 1. Perceived effectiveness of various types of signing.

Signing of TWLTL	Number of Respondents	Percentage Perceiving Service Improvement			Percentage Perceiving Safety Improvement		
		Significant	Slight	None	Significant	Slight	None
No signing	10	50	40	10	50	30	20
Overhead	14	86	14	0	57	29	14
Side mounted	5	80	20	0	80	20	0
Painted arrow	10	70	20	10	70	20	10
Composite, one device only	29	72	24	4	66	24	10
Arrows and side mounted	10	80	20	0	67	33	0
Overheads and side mounted	4	67	33	0	50	50	0
Arrows and overhead	8	100	0	0	87	13	0
Composite, two devices	22	86	14	0	71	29	0
All three signing devices	5	100	0	0	80	20	0

Little effort has been made to measure left-turn demand or to establish standard values or ranges of values that would specifically dictate conditions for installation of a TWLTL.

Traffic Volume

In the literature, successful TWLTL operations were described as widely ranging traffic volumes [8000 to 31 000 average daily traffic (ADT)], and traffic volume was not identified as a particularly critical factor except when it approached capacity. The references to roadways operating at or near maximum capacity (3) only predicted that the value of the TWLTL in reducing congestion under such conditions might become questionable because of the unavailability of gaps of sufficient size in the approaching traffic to allow the left-turn movement.

In such cases, however, if direct left-turn access must be provided but signalization cannot be used to alter gap size or distribution favorably so as to accommodate left-turning vehicles, then left-turn storage of some type becomes even more necessary. A policy directive of the Washington State Department of Highways (6) specifically states that the following minimum and maximum volumes should prevail: 5000 to 12 500 ADT on two-lane roadways and 10 000 to 25 000 ADT on multilane roadways.

Speed Limit

The existing speed limit on a highway facility does not appear to be a critical factor for consideration in TWLTL applications except in the general sense. The reports in the literature that refer to speed reinforce its consideration as a general warrant and refer to TWLTLs operating at speeds that range from 40.3 to 80.5 km/h (25 to 50 mph). Concern has been expressed about TWLTL operations at speeds higher than these (5, 6, 8) because of the increased accident potential and at speeds lower than these because of the possibility that impatient drivers may use the median lane to pass slower vehicles. Neither concern has been sufficiently supported by data to rule out TWLTL applicability at wider ranges of speed.

Spacing of Existing Intersections

The effects of intersection spacing on TWLTL application have not been thoroughly examined or documented in the literature. The studies that did comment on intersection spacing (2, 5) provided very general testimony about the adverse effects of closely spaced intersections without defining any specific minimum desirable limitation on spacing. Their concerns were based only on the problem of maintaining a sufficient block length to accommodate exclusive left-turn-only lanes at each intersection and also some minimum length for the TWLTL in midblock. Perhaps the major importance of intersection spacing lies in its contribution to the effect on local traffic circulation patterns and therefore on alternative access.

Economic Considerations

Only two reports that attempted a detailed economic analysis of TWLTLs were located (9, 10). These studies used a method to determine user benefits based on reductions in fatal, injury, and property-damage accidents. The first evaluation (9) determined that the TWLTL installation would pay for itself in less than 2 years; the accrued benefits for the four TWLTLs in the other study (10) were such as to surpass the improvement costs in 7 years. Significantly, all five installations involved some capital costs as a result of widening of the pavement.

Where sufficient pavement width is already available, the TWLTL installation primarily involves only restriping and signing so that in many instances the work can be accomplished by force account with maintenance personnel and equipment rather than by more costly contracting procedures.

Investigation of the economic impact of the TWLTL on adjacent properties has been minimal (3), but the value of this information in a typical traffic engineering study is limited.

Safety

The 15 reports that were reviewed for safety considerations represented accident experience at approximately 50 TWLTL installations. However, because of the great variation in the detail and the methods of the many TWLTL evaluation studies, no quantitative, composite figures for accident reduction could be derived that would be truly representative of all the TWLTLs investigated in the literature.

Only a few reports included data on fatal and personal-injury accidents or gave particular emphasis to investigating accident severity. This was a surprising omission, but the studies that did include such data offered conclusive evidence that TWLTLs significantly reduce accident severity (11). In their investigations, Sawhill and Neuzil (4) found that the TWLTL accident is somewhat less severe than the non-TWLTL accident, and the two studies by the Michigan Department of State Highways (9, 10), which represented experiences at five TWLTL installations, substantiated their findings.

The types of accidents that are acknowledged to be most commonly affected by the installation of the TWLTL, and therefore the types of accidents to which the improvement has subsequently been most directed—rear-end, sideswipe, and midblock left-turn collisions—were found either to decrease substantially in numbers or at least to have had their growth rates significantly retarded in the face of regional trends of increasing accident occurrence in nearly every case documented in the literature.

The head-on collision, which has been a major concern underlying every decision to install a TWLTL because of deadly past experience with the old median bidirectional passing lanes, has been proved in every study to be an uncommon occurrence and of negligible concern (12).

Design Details

Number of Lanes and Lane Width

The literature documents successful TWLTL operations on facilities that have one, two, or three through lanes in each direction. No data are available that favor any of the three basic configurations generally in use from an operational standpoint, but the five-lane section is the most common. In addition, there is nothing to prevent the TWLTL from being used in applications where there is unbalanced distribution of lanes, but this configuration has not been documented in the literature.

The only conclusive value of lane width discernible from the literature is the 3-m (10-ft) minimum width, which appears to be universally accepted. Until such time as optimum lane widths are defined and uniformity is obtained through strict adherence to the MUTCD, practical experience will dictate that the current 3- to 4.5-m (10- to 15-ft) range of lane width continue to be used.

Signing and Pavement Marking

The review of the literature points out that current signing and pavement-marking practices are still best characterized by a considerable lack of uniformity (13, 14, 15). One point worth noticing is that BEGIN TWO-WAY LEFT-TURN LANE signs in medial island areas were subject to repeated damage unless they were placed with proper clearance (13). Standards for signing and marking TWLTLs have been developed and included in the MUTCD (18).

Treatment at Intersections

The standard method of pavement marking in the MUTCD provides separate left-turn bays at major intersections while permitting the TWLTLs to be carried up to minor intersections (18). This solution seems logical, but the literature reviewed did not provide formal evidence either for or against this practice.

Evaluation

Accident Characteristics

The conflict study used in our field studies can provide immediate feedback after the installation of TWLTLs and thus would be more useful than the before-and-after accident studies reported in the literature. Accident patterns take a considerable amount of time to develop.

Proper Use of TWLTLs

Since TWLTLs are still unknown in many cities, a certain segment of the driving population is not familiar with them. Two-way traffic in a lane is foreign to normal driving instincts. The literature and our field studies indicated that improper use could be a problem, at least initially. Improper use can only be prevented by educating the public before installation of the lanes. Some extensive and equally effective approaches have been reported in the literature (9, 10). Deliberate violation of the rules, such as driving in the TWLTL for an excessive distance, can only be eliminated by enforcement.

FIELD STUDIES

Purpose

Before-and-after studies were completed at three sites in Ohio where the introduction of TWLTLs was not accompanied by other major improvements. The purpose was to measure the effect of TWLTLs on traffic flow conditions and on safety.

Data Collection

Data on travel time and delay were collected by using a vehicle equipped with a tachograph. Through volumes were counted by mechanical recorders, and turning volumes were tallied by visual observation. Data on traffic conflicts were collected by a team of specialists from the Ohio Department of Transportation.

Running speeds were computed by eliminating from the travel time those delays that were in no way related to midblock left turns. Average running speeds were computed from approximately 40 runs, usually made between 9:00 a.m. and 6:00 p.m. on two weekdays and on one Saturday, for each phase—before, immediately after, and 6 months after the installation of TWLTLs.

Only running speeds and conflicts are presented here. Details of the field studies are given elsewhere (1).

Site 1, Painesville, US-20

Characteristics

Site 1 had the following characteristics: length—1.5 km (0.95 mile), width (used as four-lane roadway although centerline only was marked)—10.9 m (36 ft), volume—16 320 ADT, speed limit—72.5 km/h (45 mph) posted, and adjacent land use—commercial strip development.

Reconstruction

This four-lane arterial was restriped as a three-lane roadway. The TWLTL was identified by overhead signs and pavement arrows.

Effect on Flow

Average running speeds and directional hourly traffic volumes are given below (1 km/h = 0.62 mph):

Direction	Period	Average Speed (km/h)	Hourly Volume
Eastbound	Before installation	55.47	405
	After installation	49.71	401
Westbound	Before installation	53.45	508
	After installation	45.81	574

The elimination of one through lane in each direction offsets the beneficial effects of the TWLTL.

Effect on Safety

Brake applications were reduced 22 percent, from 614 to 480, but weavings increased 78 percent, from 105 to 187. The increase in weavings prompted us to investigate driver behavior further. A time-lapse film recorded 548 left turns, with the following results:

1. Eighteen (or 3 percent) did not use the TWLTL at all.
2. Thirty-two (or 6 percent) turned into the TWLTL at an angle, and part of the vehicles protruded into the through lanes.
3. Seventy-eight (or 14 percent) moved into the TWLTL only partially, and the two right-hand wheels remained in the through lanes. This type of improper use might have been caused by the old centerline, which was not properly removed. (A similar problem was observed at site 2, and proper removal of the line eliminated the problem.)

Results

The conversion of two through lanes into a TWLTL improved the access function of the roadway at the expense of the movement function. During the short peak periods, the impact was much worse than the above-average speeds would indicate. The traffic backup in the area prompted some impatient drivers to use the TWLTL as a passing lane. The obvious solution would be to operate the median lane as a TWLTL during off-peak periods and use it as a reversible flow lane during peak periods. Several such installations are now in operation in some cities.

Site 2, Cincinnati, OH-264

Characteristics

Site 2 had the following characteristics: length—1.48 km

(0.92 mile), width (four lanes with different types of medians on some parts)—17.9 m (59 ft), volume—17 610 ADT, and adjacent land use—commercial strip development.

Reconstruction

This four-lane roadway was restriped as a five-lane roadway. Overhead signs and pavement arrows were used to identify the TWLTL.

Effect on Flow

Running speeds were obtained before, after, and 6 months after installation of the TWLTL. Although speeds increased slightly, the increase is not statistically significant (1 km/h = 0.62 mph):

Direction	Period	Average Speed (km/h)	Hourly Volume
Eastbound	Before installation	51.97	762
	After installation	54.38	798
	6 months after installation	53.71	745
Westbound	Before installation	54.02	886
	After installation	56.32	887
	6 months after installation	55.39	727

Running speeds were quite satisfactory during the before period, and thus the possibilities for improvement were limited.

Effect on Safety

Brakings and weavings are given below:

Period	Number of Brakings	Number of Weavings
Before installation	575	589
After installation	685	530
6 months after installation	485	565

There was considerable variation in conflicts at different sections of the roadway. During the first data collection after installation of the lane, it was quite obvious that many drivers did not know how to use the TWLTL properly. Consequently, three samples of driver behavior—totaling 668 left turns—were observed:

1. Forty-seven (or 7 percent) did not use the TWLTL at all.
2. Seventy (or 10.5 percent) turned into the TWLTL at an angle, protruding into the path of through traffic.
3. One hundred and twenty-six (or 18.9 percent) weaved into the TWLTL only partially. This type of behavior was especially frequent in those areas where the old centerline had not been properly removed.

This high frequency of improper use caused conflicts in through lanes.

The centerline was eventually properly removed. A similar observation was scheduled for the study 6 months after installation to check for improvement after a learning period. By this time, however, improper left turns were so infrequent that data collection was discontinued.

Results

The results of the field studies do not indicate a drastic improvement in running speeds and conflicts. Traffic conditions were already quite satisfactory during the before period. The advantages of the TWLTL will become more obvious when traffic volumes increase in the area.

Site 3, Mansfield, US-42

Characteristics

Site 3 had the following characteristics: length—1.3 km (0.8 mile), width (two lanes)—9.4 m (31 ft), volume—14 070 ADT on northern half and 12 940 ADT on southern half, speed limit—56.4 km/h (35 mph) on northern half and 72.5 km/h (45 mph) on southern half, and adjacent land use—commercial (more intensive on northern half).

Reconstruction

By improving a narrow strip of the shoulder, this roadway was widened to 10.9 m (36 ft). The widening reduced the shoulder to less than 1 m (3.3 ft) on the northern half. The through lanes were reduced in width from 4.5 to 3.5 m (15 to 11.5 ft). The TWLTL is 3.9 m (13 ft) wide.

Effect on Flow

Running speeds and directional volumes are given below for the two sections separately (1 km/h = 0.62 mph):

Section	Period	Average Speed (km/h)	Hourly Volume
North			
Northbound	Before installation	56.68	490
	After installation	59.10	491
	6 months after installation	62.16	393
Southbound	Before installation	62.00	306
	After installation	64.09	295
	6 months after installation	64.00	330
South			
Northbound	Before installation	47.18	473
	After installation	48.31	527
	6 months after installation	52.53	NA
Southbound	Before installation	48.47	521
	After installation	49.11	416
	6 months after installation	52.88	568

In spite of the reduced lane width, there was a small, statistically significant increase in running speed.

Effect on Safety

Braking and weaving conflicts are summarized below:

Period	Number of Brakings	Number of Weavings
Before installation	1327	245
After installation	567	22
6 months after installation	833	48

The reduction in conflicts is dramatic. The difference between the after and 6-months-after time periods cannot be explained by the available data.

Results

The introduction of the TWLTL even at the expense of narrowing both through lanes resulted in a measurable improvement in traffic flow and safety.

IMPLEMENTATION GUIDELINES FOR TWLTLs

Implementation guidelines have been developed for traffic engineers who have had little or no experience with TWLTLs. A step-by-step decision-making process is outlined, but the traffic engineer must apply engineering judgment every step of the way.

The initial step involves the documentation of existing conditions so that the problem can be properly defined. By extending the principle of providing separate storage lanes for left-turning vehicles at intersections, TWLTLs are intended to shadow midblock left-turning vehicles from through traffic. Consequently, the objectives of the review of existing conditions are to establish that

1. A conflict between midblock left turns and through traffic exists and
2. The particular solution offered by the TWLTL is both potentially feasible and desirable.

To this effect, information is needed in three areas: existing physical conditions (both transportation and land use), existing traffic conditions, and accident histories. The following series of relevant items provides a checklist type of approach to the review of existing conditions.

Establish Conflict

Physical Conditions

1. Driveway spacing—Identify the spots where conflicts may occur.
2. Type and intensity of land use—Identify access needs, which determine the frequency and time distribution of conflicts.
3. Level of development—Establish the stability of current access needs.

Strip commercial developments and, to a lesser degree, multiple-unit residential areas generate traffic throughout the day. Industrial areas tend to generate morning and evening peak traffic.

Existing Traffic Conditions

1. Traffic volumes—The combination of through volumes and turning volumes gives a measure of the potential conflict on a given road section.
2. Flow characteristics—Directional distribution and peaking characteristics of both through and turning traffic give a more accurate indication of the conflict. Some measurement of congestion will indicate the level of the problem, which may have been caused (mostly or partially) by midblock turns.

Engineering judgment is needed to interpret and evaluate this information. Since the level of the conflict at any driveway is a complex function of opposing volumes, left-turning volumes, and through traffic and the level of the conflict on a roadway segment is a function of the conflicts at all the driveways, the establishment of quantitative guidelines was not attempted.

Accident History

Midblock sideswipe and rear-end accidents are typical results of conflict between delayed left-turning vehicles and through vehicles (weaving to avoid entrapment or braking to stop in the through lane behind a turning vehicle). The interpretation of these data will probably require a comparison of accident rates with accident experience on other arterials that carry similar volumes without midblock access conflicts.

Establish Appropriateness of TWLTL

Physical Conditions

1. Driveway spacing—Provide a basis for comparison of TWLTL with channelized left turns or other alternatives. Closely spaced driveways indicate a potential for TWLTLs.
2. Type and intensity of land use—Activities that generate left turns throughout the day will probably stimulate the development of remaining undeveloped lots. The provision of a raised median would have the opposite effect or attract only those establishments that do not generate much traffic.
3. Ease of alternative access—The conditions must be evaluated so that the relative attractiveness of TWLTLs can be evaluated in relation to alternative techniques of access control.
4. Distance between intersections—Since intersections often require channelized left-turn storage lanes, a very short block would not be appropriate for TWLTLs.
5. Section length—In urban areas, where TWLTLs are common, even extremely short TWLTLs work satisfactorily. Pioneering efforts, however, should concentrate on longer sections, probably several blocks long.
6. Number of lanes—Three- and five-lane applications are common. Some existing seven-lane installations have had accident records, and others have been reported to work well.
7. Pavement width—The TWLTL should be at least as wide as left-turn lanes. Lanes wider than 4.8 m (16 ft) might encourage two-lane use. If no excess pavement width is available, pavement widening will add to installation costs.
8. Right-of-way limits—Since TWLTLs improve access to adjacent properties, property owners may tend to cooperate when expansion of the right-of-way is needed for this purpose.
9. Curb parking—Eliminating curb parking is often the most convenient way to obtain the needed extra pavement width.
10. Sight distance—On higher speed roadways (especially in semirural areas), the provision of sufficient sight distance may require special attention.
11. Speed limit—Speed limits may need to be reevaluated (TWLTLs are reported to work in all speed ranges).

By reserving one traffic lane for left turns only, TWLTLs reduce the conflict between midblock left turns and through traffic. The source of this extra lane width requires careful consideration. Pavement widening increases the initial investment, elimination of curb parking reduces accessibility, and reduction or narrowing of through traffic lanes affects the through capacity of the roadway.

Physical conditions may require the reconsideration of the proper function of a given arterial and reduce through left-turn conflict by limiting either through or turning traffic. Signalization of major driveways, prohibition of some left turns, and provision of access from side streets are some examples of alternative approaches.

Existing Traffic Conditions

1. Traffic volumes—Existing through volumes and the capacity of major intersections should be investigated to determine the through capacity requirements of the midblock area. This must be a major factor when the source of the required pavement is considered.

2. Flow characteristics—Distribution of through and turning traffic volume during the day may be an important consideration.

The time-related distribution of turning traffic during the day is a function of the use of adjacent lanes. The center lane could be operated as a reversible-flow lane in conditions of peak-hour through traffic and as a TWLTL during off-peak hours. It would be advisable, however, to reserve this type of application for urban areas where TWLTLs have been accepted and extensively used.

Accident History

Since TWLTLs remove left-turning vehicles from through lanes, they are effective in reducing rear-end accidents. When TWLTLs are properly used, left-turning vehicles are completely shadowed from through traffic. In addition to protecting vehicles as they prepare to enter a driveway, a TWLTL provides a refuge for left turns made from driveways.

Future Development

Before the final selection of access-control needs, some attention must be paid to future conditions in terms of both access needs and volumes of through traffic.

1. Access needs—The selection of methods of access control will influence the future, especially on arterials where adjacent land development has not yet been stabilized. Increased accessibility stimulates land development. If, for example, the future land-use goals of a community include containment of strip commercial development, the TWLTL is not the best choice. A restrictive median that concentrates and controls access points might be a more logical choice.
2. Through traffic needs—The TWLTL has some potential for increasing the carrying capacity of arterials beyond the obvious improvement provided by the separation of midblock left-turning vehicles. Some examples of reversible lane operation during peak hours and even a reversible lane and separate bus lane combination have been reported in the literature (16, 17). The increasing acceptance of TWLTLs will eventually make it feasible to take advantage of bolder variations of this sound concept.

In addition, it must be remembered that TWLTLs provide such emergency service as a detour lane during construction, a detour lane during blocking of the through lane by vehicle breakdowns, and a path for emergency vehicles during congested periods.

ACKNOWLEDGMENTS

This project was sponsored by the Ohio Department of Transportation and the Federal Highway Administration, U.S. Department of Transportation. The contents do not necessarily reflect the official views or policies of the Ohio Department of Transportation or the Federal Highway Administration.

I wish to acknowledge the major contributions made by the following graduate research associates: Thomas Lunt (literature review), David Morena (questionnaire survey), and Nagui Rouphail and Fouad Bayomy (field studies). Richard D. Paddock of the Ohio Department of Transportation provided help in the conflict studies.

Special thanks are due to all transportation engineers in the country who cooperated with us in the opinion survey.

REFERENCES

1. Z. A. Nemeth. Development of Guidelines for the Application of Continuous Two-Way Left-Turn Median Lanes. Ohio State Univ., Final Rept. EES470, July 1976.
2. Report on a Study of Two-Way Left Turn Lanes. Georgia Division, Southern Section, ITE, Nov. 1975.
3. The Alexis Road Two-Way Left Turn Lane. Division of Traffic Engineering, Toledo Department of Public Safety, June 14, 1974.
4. R. B. Sawhill and D. R. Neuzil. Accident and Operational Characteristics on Arterials Utilizing Two-Way Median Left-Turn Lanes. Transportation Research Group, Univ. of Washington, Traffic and Operations Series, Research Rept. 7, Oct. 1962.
5. R. B. Sawhill and J. W. Hall. Investigation of Left Turn Movements on Arterial Streets and Highways. Transportation Research Group, Univ. of Washington, Traffic and Operations Series, Research Rept. 13, Nov. 1968.
6. Two-Way Left Turn Lanes. Washington State Department of Highways, Seattle, Policy Directive D 24-15 (HT), Sept. 13, 1973.
7. Report to the Transportation Committee: Two-Way Left Turn Lanes. Toronto Department of Roads and Traffic, Sept. 29, 1970.
8. Report of the Subcommittee on Continuous Left Turn Lanes. Uniformity Committee, Western Section, ITE, May 18, 1964.
9. Evaluation of a Safety Project. Traffic and Safety Division, Michigan Department of State Highways, Lansing, Dec. 1968.
10. Evaluation of Four Safety Projects: Widening a Four-Lane Roadway to Five Lanes and Providing a Center Lane for Left Turns. Traffic and Safety Division, Michigan Department of State Highways, Lansing, 1972.
11. An Operational Review: Two-Way Left Turn Lanes, Four-Way Stop Control, Stop Versus Yield Control. Fort Lauderdale Traffic Engineering Department, Rept. 71-04, July 1971, pp. 1-8, 22-26.
12. Summary—Before and After Study of Two-Way Left Turn Lanes. San Diego Transportation and Traffic Division, Nov. 1969.
13. Two-Way Left Turn Observances and Accident Study. Costa Mesa Traffic Engineering Department, CA, June 1968.
14. B. Page. New Five-Lane Highway Section Opens. District 2, Texas Highway Department, Fort Worth, News Service Release, April 1974.
15. Proposed Two-Way Left Turn Lanes Along Burnet Road. Austin Traffic and Transportation Department, TX, Aug. 1968.
16. Austin's Reversible Lane System—A Dynamic Traffic Operation Technique. Winter Meeting, Texas Section, ITE, Jan. 1974.
17. Proposed Changeable Message Signing System for South Connecticut Street. Seattle Department of Engineering, 1974.
18. Manual on Uniform Traffic Control Devices for Streets and Highways. Federal Highway Administration, U.S. Department of Transportation, 1971.

Publication of this paper sponsored by Committee on Operational Effects of Geometrics.

Selection of Median Treatments for Existing Arterial Highways

Douglas W. Harwood, Midwest Research Institute, Kansas City, Missouri
John C. Glennon, Transportation Consulting Engineer, Overland Park, Kansas

Median treatments are an important means of reducing accidents and delay on urban arterial highways. Five common median treatments are (a) two-way left-turn lane, (b) continuous left-turn lane, (c) alternating left-turn lane, (d) raised median divider with left-turn deceleration lanes, and (e) median barrier with no direct left-turn access. A benefit-cost comparison of these treatments that considers the accident reduction, delay reduction, and construction cost for median treatments installed in existing arterial highways is reported. The analysis is based on a literature review and reasonable assumptions regarding the effectiveness of the median treatments. The result of the benefit-cost analysis is a selection guide that can be used by a designer to determine the optimal median treatment for an arterial highway based on geometric and operational conditions.

Many urban arterial highways in the United States have serious operational and safety deficiencies. These deficiencies are often the combined result of high and steadily growing traffic volumes and of a high density of driveways resulting from a lack of effective access control. These highways often have nonintersection left-turn movements that are nearly continuous in space and time. If unrestrained, these demands can result in both high accident rates and large delays to through motorists.

Traffic engineers responsible for arterial highways have long recognized the important role of median treatments in alleviating the operational and safety deficiencies described above. Indeed, many of the common safety and operational problems are amenable to solution in no other way. Left-turning vehicles are often the cause of accidents and delays to through vehicles, and only median treatments can alleviate these left-turn problems. Such problems are often continuous on long stretches of arterial highway, and only the continuous solution provided by a median treatment is practical.

Five basic median treatments have the potential to improve traffic operations and safety for continuous sections of existing arterial highways. These are (a) the two-way left-turn lane (TWLTL), (b) the continuous left-turn lane (CLTL), (c) the alternating left-turn lane (ALTL), (d) the raised median divider (RMD) with left-turn deceleration lanes, and (e) the median barrier (MB) with no direct left-turn access. The design and operational characteristics of these treatments are briefly described in the following section. Most of these treatments are currently used by at least some agencies, but the traffic engineer needs a rational basis for selecting a median treatment that is both cost-effective and operationally appropriate for a given highway section. The discussion in this paper provides the framework for a rational method of selecting appropriate median treatments.

DESIGN AND OPERATIONAL CHARACTERISTICS

The five median treatments fall into two distinct categories. The first three treatments use median lanes that do not physically restrict the movement of traffic across the median. The last two techniques use raised medians that limit crossings to those openings selected by the designer. The design and operational characteristics of the five median treatments are discussed below. More detailed descriptions of these techniques can be found in recent reports by Azzeh and others (1) and Glennon and others (2).

Two-Way Left-Turn Lane

The standard design for a two-way left-turn lane specified by the Manual on Uniform Traffic Control Devices (MUTCD) (3, Figure 3-4a) is shown in Figure 1. The major design requirement for this technique is the median width, which should be at least 4.2 m (14 ft).

A two-way left-turn lane is intended to remove left-turning vehicles from the through lanes and store those vehicles in a median area until an acceptable gap in opposing traffic appears. The two-way left-turn lane completely shadows turning vehicles from both through-lane traffic streams. Thus, reductions in the severity and frequency of accidents will result. Frequency is reduced by removing stopped or slow left-turning vehicles from the through lanes, and severity is reduced by allowing additional perception time to reduce left-turn crossing conflicts. Delay to through vehicles is also reduced because left-turning vehicles and queues do not block the through lanes.

The two-way left-turn lane is operationally warranted on arterial highways that have average daily traffic (ADT) volumes higher than 10 000 and traffic speeds faster than 48 km/h (30 mph). The number of driveways should exceed 60 in 1.6 km (1 mile), and there should be fewer than 10 high-volume driveways. Left-turn driveway maneuvers in 1.6 km should total at least 20 percent of the through traffic volume during peak periods. High rates of accidents that involve left-turn maneuvers can also warrant this technique.

Continuous Left-Turn Lane

The standard design for a continuous left-turn lane is shown in Figure 2 (based on MUTCD Figure 3-4b). This technique is similar to the two-way left-turn lane except that it provides individual left-turn lanes for each direction of traffic. Each left-turn lane is continuous except that far-side channelizing islands are placed to prevent through movements at signalized intersections. Left-turn vehicles can be stored in the continuous left-turn lane until an acceptable gap in opposing traffic appears. The continuous left-turn lane completely shadows turning traffic from both traffic streams. Accident frequency is reduced by removing stopped or slow vehicles from the through lanes, and accident severity is reduced by allowing through vehicles additional perception time to avoid left-turn crossing conflicts. Delay to through vehicles is also reduced because left-turn vehicles and queues do not block the through lanes.

The major design difference between this technique and the two-way left-turn lane is the required median width. A 7.2-m (24-ft) wide median is needed for this technique. This width will accommodate two 3.6-m (12-ft) turning lanes. At locations where 7.2 m (24 ft) is not available for median width, it is advisable that a two-way left-turn lane be considered. Since the turning lanes are continuous, this technique should be applied over sections at least 0.4 km (0.25 mile) in length.

Alternating Left-Turn Lane

The design of an alternating left-turn lane is shown in Figure 3. The alternating left-turn lane will allow one traffic direction to have the opportunity to cross the median into driveways and, after a specified distance, the left-turn lane is physically opened to the opposing direction of traffic. Thus, both the directions have a unique left-turn lane available for continuous left-turn maneuvers over a limited section of highway. Left-turn access to some driveways is prevented because, when the left-turn lane is available to one traffic direction, the opposing traffic cannot attempt a left turn.

The striping scheme shown in Figure 3 is not readily recognized by today's motorist as delineating a left-turn lane. No striping criteria have been universally adopted for use with a technique such as this. The use of turn arrows should help to reduce confusion.

An important design consideration for the alternating left-turn lane is the configuration of the deceleration taper. In this technique, the deceleration taper not only delineates the correct deceleration path but also serves to separate the left-turn lane for different traffic directions.

Reductions in the frequency and severity of accidents will result from the implementation of this technique. Frequency is reduced by removing stopped or slow-moving vehicles and queues from the through lanes, and severity is reduced by allowing through vehicles additional perception time to avoid left-turn crossing conflicts. Delay to through vehicles will also be reduced because left-turning vehicles will not block the through lanes.

The major advantage of implementing this technique instead of other median treatments lies in the minimum median width required to accommodate the left-turn lane. Since only one lane is used in the median for left-turn movements, the width of the median should be as wide as the turning lane itself. Whereas other treatments require 4.2- to 7.2-m (14- to 24-ft) medians for left-turn movements, this treatment requires only a 3.6-m (12-ft) median. The value of this treatment for application on narrow-median highways is most evident at locations where pavement widening or right-of-way acquisition would be required for the wider medians.

Raised Median Divider With Left-Turn Deceleration Lanes

The raised median divider with left-turn deceleration lanes, shown in Figure 4, promotes safety and through-traffic service by preventing left turns and U-turns across the median except at a few designated locations. Access is provided by left-turn lanes at intersections and major driveways. In addition to preventing left turns at minor driveways, the raised median divider reduces friction in the traffic stream by separating opposing traffic.

This technique reduces the frequency of total conflicts by reducing the number of basic conflict points at all minor driveways. More important, it completely eliminates the more hazardous points of crossing conflict at these driveways. For intersections and major driveways, the frequency and severity of conflicts

Figure 1. Two-way left-turn lane.

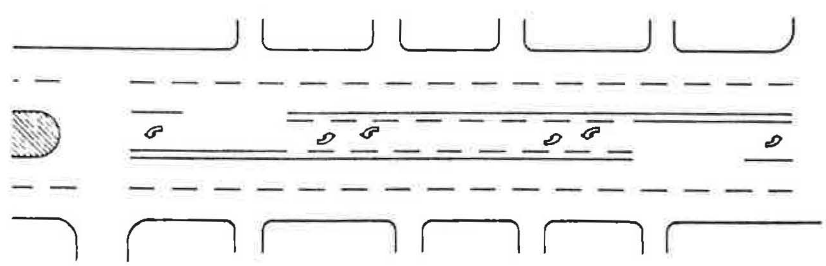


Figure 2. Continuous left-turn lane.

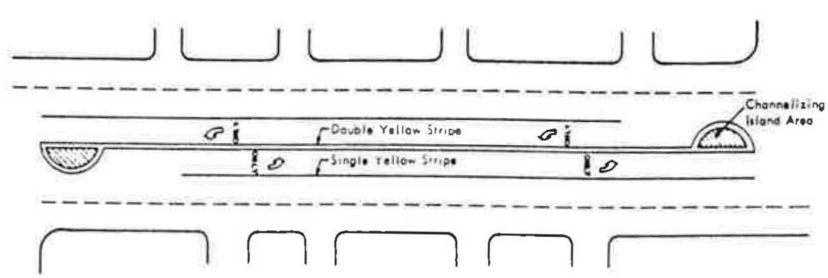
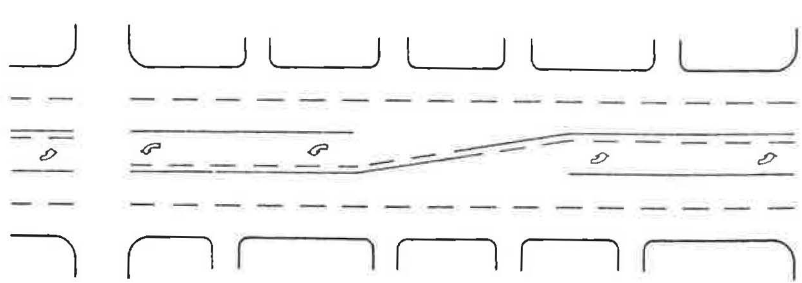


Figure 3. Alternating left-turn lane.



associated with left-turn vehicles are reduced by allowing deceleration and shadowing of these vehicles in left-turn lanes.

The median divider usually reduces the total number of driveway maneuvers. However, the maximum reduction in the frequency of conflicts is moderated by increases in right-turn volumes at minor driveways where desired left turns are accomplished through indirect, circuitous paths.

The construction of a raised median divider often requires widening of the existing roadway. Where insufficient right-of-way has been dedicated, additional right-of-way will need to be purchased. The minimum required roadway width is 16.8 m (56 ft). This width accommodates four 3.3-m (11-ft) through lanes and a 3.6-m (12-ft) median. A more desirable design allows four 3.6-m (12-ft) through lanes and a 4.8-m (16-ft) median for a total roadway width of 19.2 m (64 ft) (Figure 4).

The most important design element for the raised median divider is the median width, which must be adequate to completely shadow left-turning vehicles from through vehicles. The desirable minimum median width is 4.2 m (14 ft). This width provides a 3.6-m (12-ft) deceleration storage lane and a 0.6-m (2-ft) raised median at median openings. However, a 4.8-m (16-ft) median width is recommended, and a 6.6-m (22-ft) width is required if U-turns are permitted.

The required minimum deceleration length is that distance required if a vehicle is to make a comfortable stop from the average running speed on the highway. The storage length should be sufficient to store the maximum expected vehicle queue. As a minimum, storage length for at least two passenger automobiles should be provided. The spacing of median openings is dictated by the length of the deceleration lane, which varies from 90 to 300 m (300 to 1000 ft) for design speeds from 48 to 72 km/h (30 to 45 mph).

Median Barrier With No Direct Left-Turn Access

The final median treatment considered here is the median barrier with no direct left-turn access. This design has no left-turn deceleration lanes, but instead left turns are accomplished by means of indirect left-turn ramps—cloverleaf loops or jughandles—at median openings. Figures 5 and 6 show these two basic designs. The cloverleaf design (Figure 5) is recommended when the distance between major driveways or intersections is less than 1.6 km (1 mile). The jughandle design (Figure 6) is recommended when major driveways or intersections are spaced at 1.6 km or more. This treatment incorporates a New Jersey type of barrier or a simple barrier curb in the median and eliminates all direct left turns and U-turns along the highway.

The median barrier with no direct left-turn access reduces the number of basic conflict points and totally eliminates the more hazardous crossing conflicts at driveways in much the same way the raised median divider does. Furthermore, the frequency of rear-end conflicts on the through lanes is expected to decrease as a result of the elimination of direct left turns. On the other hand, the frequency of right-turn conflicts at minor driveways will probably increase in proportion to the number of indirect left turns. The reduction in points of crossing conflict at driveways is partially offset by the creation of additional basic conflict points at indirect left-turn locations. However, this trade-off is minimized if these locations are signalized.

A much narrower median is required for this treat-

Figure 4. Raised median divider with left-turn deceleration lanes.

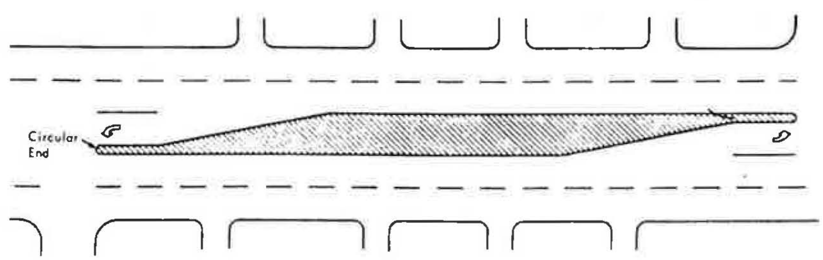


Figure 5. Median barrier with indirect left-turn ramp (cloverleaf loop).

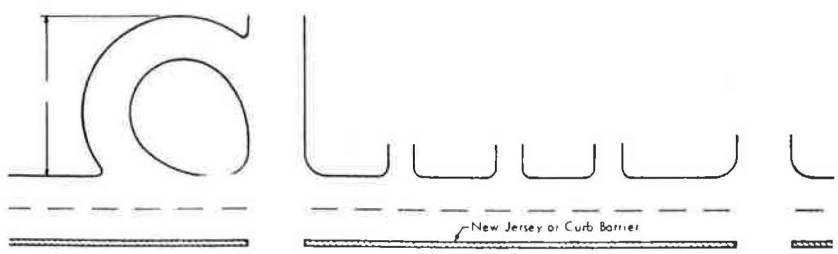
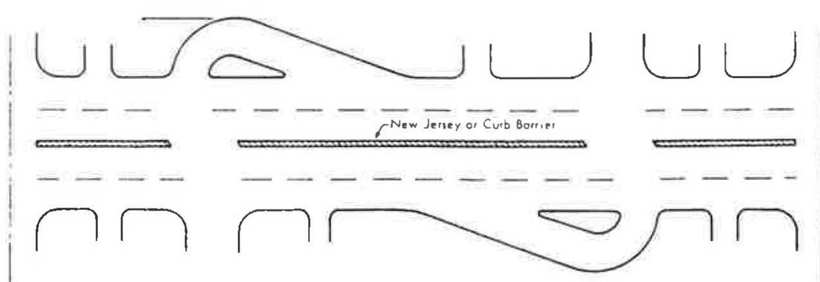


Figure 6. Median barrier with indirect left-turn ramp (jughandle).



ment than for the raised median divider because this treatment eliminates the left-turn deceleration lanes. The desirable median width for the barrier is 1.8 m (6 ft), which is sufficient to accommodate a 0.6-m (2-ft) wide barrier with a 0.6-m (2-ft) clearance on each side. However, a right-of-way width of more than 45 m (150 ft) is needed at the jughandle or cloverleaf sites, and this requirement alone may render this design impractical at many locations.

EFFECTIVENESS OF MEDIAN TREATMENTS

The selection of the optimal median treatment should be based on its effectiveness in reducing both accidents and delay. Unfortunately, the effectiveness of most median treatments has not been evaluated consistently. This lack of proven evaluations does not eliminate the need to make rational choices among the available median treatments and should not deter the use of the best available information to compare alternatives. Estimates of effectiveness can be developed from the available literature and reasonable assumptions. The effectiveness of the five median treatments in reducing both accidents and delay is considered below.

Accident Reduction

The effectiveness of median treatments in reducing accidents can be estimated by (a) estimating the accident experience of typical arterial highways, commercial driveways, and signalized intersections; (b) estimating the number of driveways and intersections per kilometer on typical arterial highways; and (c) determining the number of accidents per kilometer per year that would be reduced by each median treatment. Level of development and highway, driveway, and crossroad ADT—measurements used in the tables throughout this paper—are defined below:

Item	Definition
Level of development	
Low	< 30 driveways
Medium	30-60 driveways
High	> 60 driveways
Highway ADT	
Low	< 5000 vehicles/d
Medium	5 000-15 000 vehicles/d
High	> 15 000 vehicles/d
Driveway ADT	
Low	< 500 vehicles/d
Medium	500-1500 vehicles/d
High	> 1500 vehicles/d
Crossroad ADT	
Low	< 500 vehicles/d
Medium	500-1500 vehicles/d
High	> 1500 vehicles/d

Table 1 gives expected annual accident frequency for a 1.6-km (1-mile) section of arterial highway for three ADT levels and three levels of development. The values are based on regression equations developed by Mulinazzi and Michael (4). The derivation of these values from the Mulinazzi and Michael regression equations is documented by Azzeh and others (1).

Annual accident frequencies for typical commercial driveways are given in Table 2. A prediction equation for the expected accident experience of four-way, unsignalized intersections on divided highways was developed in a study by McDonald (5). A commercial driveway is essentially a three-way intersection that has only 9 conflict points whereas a four-way intersection has 32 conflict points. The accident predictions in Table 2 were obtained by multiplying the accident frequencies from McDonald's equations by $^9/_{32}$. Accidents probably do not correspond directly with the numbers of conflict points, and some particular maneuvers definitely have a higher frequency of conflicts under certain conditions. However, the procedure is valuable in making comparisons.

Annual accident frequencies for four-way signalized intersections on arterial highways are given in Table 3. These data are based on the work of Webb (6).

For evaluation purposes, 1.6 km (1 mile) of a typical arterial highway is assumed to have two signalized intersections. The assumed distribution of driveways per 1.6 km for each level of development is given below:

Level of Development	Number of Driveways by Driveway Volume		
	High	Medium	Low
High	0	10	65
Medium	2	8	35
Low	2	5	8

Two-way left-turn lanes have been evaluated by several agencies: Two studies have been conducted in Michigan (7, 8), one in Sacramento, California (9), one in Seattle (10), and one by a technical council of the Institute of Transportation Engineers (11). The results of a before-and-after study of 11.6 km (6.6 miles) of two-way left-turn lanes in Michigan (12) are given below:

Type of Accident	Number of Accidents		Change (%)
	Before	After	
Left turn	94	52	-45
Rear end	238	90	-62
Right angle	92	105	+14
Sideswipe	42	39	-7
Other	66	70	+6
Total	532	356	-33

Table 1. Annual accident frequency for 1.6 km (1 mile) of typical arterial highway.

Level of Development	Accidents for Three Levels of Highway ADT		
	Low	Medium	High
Low	12.6	25.1	37.9
Medium	20.2	39.7	59.8
High	27.7	54.4	81.7

Table 2. Annual accident frequency for a typical commercial driveway.

Driveway ADT	Accidents for Three Levels of Highway ADT		
	Low	Medium	High
Low	0.26	0.45	0.62
Medium	0.63	1.1	1.5
High	0.97	1.7	2.3

Table 3. Annual accident frequency for a typical four-way, signalized intersection.

Crossroad ADT	Accidents for Three Levels of Highway ADT		
	Low	Medium	High
Low	1.1	1.9	2.6
Medium	1.9	3.2	4.4
High	2.5	4.2	5.8

This table illustrates the effectiveness of this treatment in reducing left-turn and rear-end accidents. Although increases in head-on accidents might be expected in the median lane because of conflict with opposing vehicles, the literature discounts such occurrences as infrequent. Based on all of the studies identified above, the total number of accidents on an arterial street can be expected to decrease by 35 percent after installation of a two-way left-turn lane on a four-lane arterial street. The expected accident reductions for 1.6 km (1 mile) of an arterial street have been calculated from Table 1 and are given in Table 4.

No operational studies on the continuous left-turn lane were found in the literature. Therefore, the two-way left-turn lane was used as the basis for comparison, and two operational differences between the two-way left-turn lane and the continuous left-turn lane were considered. First, the continuous left-turn lane has a separate left-turn lane for each direction of travel, which should reduce some conflicts that result from opposing vehicles using the same lane. On the other hand, motorists turning left from the continuous left-turn lane must cross the left-turn lane for the opposite direction, which therefore increases the conflict area. It seems reasonable to assume that these two effects cancel one another and therefore that the effectiveness for the continuous left-turn lane is the same as that for the two-way left-turn lane (Table 4).

Alternating left-turn lanes operate somewhat differently from two-way and continuous left-turn lanes. In addition to reducing rear-end and left-turn conflicts, alternating left-turn lanes may reduce the frequency of left-turn maneuvers by discouraging left-turn access at driveways where there are opposing left-turn deceleration lanes. One study (13) indicates a 28 percent decrease in accidents as a result of converting a section of highway to alternating left-turn operation. To completely evaluate the effectiveness of this treatment, several assumptions were made about its operational characteristics. Included in these assumptions is that left-turn access will be provided to all medium- and high-volume driveways but to only half of the low-volume driveways. The installation of left-turn lanes for these driveways is assumed to reduce their accident experience by 50 percent. For half of the low-volume driveways where the left-turn access is denied, accident experience is assumed to be reduced by 60 percent. Finally, it is assumed that the installation of the median lane makes it possible to install left-turn lanes at two signalized intersections and that the accident experience at these intersections is reduced by 50 percent. Detailed explanation and justification of these assumptions are provided by Azzeh and others (1). The resulting estimates of accident reduction for this treatment are given in Table 4.

The raised median divider is evaluated by using assumptions similar to those made for the alternating left-turn lane, but this treatment is even more restrictive operationally. The treatment is assumed to prevent left turns at all low-volume driveways and to result in a 60 percent accident reduction at these driveways. It is also assumed that left-turn deceleration lanes are installed at all medium- and high-volume driveways and at signalized intersections. This results in a 50 percent decrease in accidents at these locations. The overall accident reduction for the installation of a raised median divider is given in Table 4.

The median barrier with no direct left-turn access is similarly evaluated. The barrier is assumed to eliminate left turns at all driveways. Although this results in a 50 percent reduction in driveway accidents, an accompanying increase in accidents is associated with the two signalized, indirect left-turn locations. The net accident reduction for installation of a median barrier is given in Table 4.

Table 4. Annual reduction in accidents from median treatments for 1.6 km (1 mile) of typical arterial highway.

Level of Development	ADT	Annual Number of Accidents Reduced				
		TWLTL	CLTL	ALTL	RMD	MB
Low	Low	4.4	4.4	1.7	2.2	-2.7
	Medium	8.8	8.8	3.2	4.1	-4.0
	High	13.3	13.3	5.1	6.3	-5.0
Medium	Low	7.1	7.1	3.5	5.8	1.8
	Medium	13.9	13.9	7.1	11.2	4.7
	High	20.9	20.9	11.6	17.2	8.1
High	Low	9.7	9.7	6.4	10.7	6.3
	Medium	19.0	19.0	13.3	20.7	13.6
	High	28.6	28.6	21.0	31.2	21.3

Table 5. Annual reduction in delay from median treatments for 1.6 km (1 mile) of typical arterial highway.

Level of Development	ADT	Annual Reduction in Delay (h)				
		TWLTL	CLTL	ALTL	RMD	MB
Low	Low	0	0	0	0	0
	Medium	0	0	0	0	0
	High	0	0	0	0	0
Medium	Low	0	0	0	0	0
	Medium	2 628	2 628	2 628	2 628	0
	High	6 935	6 935	6 935	6 935	0
High	Low	0	0	0	0	0
	Medium	6 059	6 059	6 059	6 059	0
	High	17 046	17 046	17 046	17 046	0

Delay Reduction

No comparative data on the effectiveness of the five median treatments in reducing delay are available. However, four of the treatments have a very similar effect on delay. The two-way left-turn lane, the continuous left-turn lane, the alternating left-turn lane, and the raised median divider with left-turn deceleration lanes all reduce delay by removing left-turning vehicles from the through lanes to a sheltered area in the median. This results in an increase of the average running speed of through traffic. The effectiveness of these treatments in reducing delay was estimated by assuming a value for this increase in average running speed. The following assumptions were made to estimate reductions in delay for typical four-lane highways:

1. Arterials with low traffic volumes or low levels of development would not experience any increase in running speed.
2. Average running speeds on arterials without median treatments are assumed to be as given below (1 km/h = 0.62 mph):

Highway ADT	Level of Development	Average Running Speed (km/h)
Medium	Medium	56
Medium	High	48
High	Medium	48
High	High	40

3. For a medium level of development, there is an increase of 8 km/h (5 mph) in average running speed during the 2 h of each day that show the highest traffic volume. These hours are assumed to include 20 percent of all through vehicles.
4. For a high level of development, there is an in-

crease of 8 km/h (5 mph) in average running speed during the 4 h of each day when traffic volume is highest. These hours are assumed to include 35 percent of all through vehicles.

The estimated effectiveness of reduction in delay that results from these assumptions is given in Table 5.

Installation of the fifth median treatment—the median barrier with no direct left-turn access—will also result in an increase in average running speed. However, this saving is probably offset by the increase in travel time for indirect left-turning vehicles and the increased delay if the indirect crossings are signalized. For evaluation purposes, these effects are assumed to be equal and offsetting so that the net reduction in delay is zero (this assumption may be unrealistic on extremely high-volume highways where median barriers may be far more desirable than suggested by the following analysis).

COST OF MEDIAN TREATMENTS

The effectiveness of median treatments should be evaluated in relation to their costs. For this reason, construction costs have been estimated for the installation of each of the five median treatments for 1.6 km (1 mile) of a typical existing arterial highway. Three construction options, presented in order of increasing cost, are considered separately for each median treatment:

1. Option 1 assumes that the existing roadway is wide enough to permit installation of the median treatment without additional widening.
2. Option 2 assumes that pavement widening is necessary but that no additional right-of-way must be acquired.
3. Option 3 assumes that both pavement widening and right-of-way acquisition are necessary to install the median treatment.

The estimated construction costs for each median treatment and construction option are given below:

Median Treatment	Cost ($) Option 1 (existing paved median)	Option 2 (pavement widening required)	Option 3 (pavement widening and right-of-way acquisition required)
Two-way left-turn lane	8 200	280 200	501 000
Continuous left-turn lane	12 800	403 200	783 600
Alternating left-turn lane	10 200	282 200	503 000
Raised median divider	97 600	369 600	590 400
Median barrier	185 200	304 000	398 800

These estimates were determined from the following unit costs (1 m = 3.3 ft, 1 m^2 = 10.76 ft^2, and 1 km = 0.62 mile):

Construction Item	Unit Cost for Construction and Overhead ($)
Pavement striping (reflective)	2.10/m
Pavement (0.3 m thick)	24.70/m^2
Driveway patchback	24.70/m^2
Curb and gutter	26.70/m
Median barrier (New Jersey type)	66.70/m
Relocation of structures (one side of roadway)	6250/km
Right-of-way acquisition	33.30/m^2

These unit costs are based on data gathered in 1975. Naturally, they are expected to rise as time passes, but the benefit-cost comparisons in the next section should still be valid since the costs of accidents and delay time are presumably rising as well. The service lives of all capital items are estimated at 20 years except for pavement striping for which the life is estimated at 2 years.

BENEFIT-COST RATIOS FOR MEDIAN TREATMENTS

The five median treatments have been compared on the basis of their benefit-cost ratios. For purposes of this study, the benefit-cost ratio (BC) is defined as

$$BC = [(AR)(AC) + (DR)(DC)]/CC\,[CRF]_i^n \qquad (1)$$

where

- AR = annual number of accidents reduced,
- AC = average cost per accident = $2800 (14),
- DR = annual hours of delay reduced,
- DC = average cost per hour of delay = $4.50 (1, 15),
- CC = total construction cost,
- $[CRF]_i^n$ = capital recovery factor at i percent for n years,
- i = minimum attractive rate of return = 7 percent, and
- n = service life = 2 years for pavement striping and 20 years for other capital items.

The benefit-cost ratio for each median treatment for each construction option is given in Table 6. Benefit-cost ratios less than 1.0 are not shown because these median treatments are not warranted under the specified conditions.

SELECTION OF MEDIAN TREATMENTS

The benefit-cost analysis presented provides a basis for selecting appropriate median treatments for arterial highways. The objective should be to select a median treatment that is not merely warranted but optimal. This objective can be accomplished by using Table 7, which summarizes the results of the benefit-cost analysis in the form of a selection guide for median treatments and construction options. The table contains a series of median treatments and construction options for each possible combination of daily traffic volume and level of development at the site under consideration. Treatment-option combinations are given in order of descending benefit-cost ratio. The optimal median treatment is the highest treatment on the list that is operationally warranted and physically feasible at the site under consideration. The width requirements of median treatments are very important in making a choice; for example, an alternating left-turn lane is preferable to a two-way left-turn lane only if construction option 1 (no widening) can be used when the two-way left-turn lane would require construction option 2 (widening).

A great many useful general conclusions about the selection of median treatments can be drawn from Table 7. For instance, at sites that have 5000 (low) ADT and driveway density of <30 in 1.6 km (1 mile), the only warranted median treatments are construction option 1 for the two-way and continuous left-turn lanes. By contrast, on highways that have >15 000 (high) ADT and driveway density of >60 in 1.6 km, all median

Table 6. Benefit-cost ratios for median treatments.

		Level of Development								
		Low			Medium			High		
Median Treatment	Construction Option	Low ADT	Medium ADT	High ADT	Low ADT	Medium ADT	High ADT	Low ADT	Medium ADT	High ADT
TWLTL	1	2.7	5.4	8.2	4.4	11.2	19.8	6.0	17.7	34.6
CLTL	1	1.7	3.5	5.3	2.8	7.2	12.7	3.8	11.4	22.1
ALTL	1	-	1.6	2.5	1.7	5.6	11.3	3.2	11.4	24.0
RMD	1	-	1.3	1.9	1.8	4.7	8.6	3.3	9.3	17.9
MB	1	-	-	-	-	-	1.3	1.0	2.2	3.4
TWLTL	2	-	-	1.2	-	1.7	3.0	-	2.7	5.2
CLTL	2	-	-	-	-	1.2	2.0	-	1.8	3.6
ALTL	2	-	-	-	-	1.0	2.0	-	2.1	4.3
RMD	2	-	-	-	-	1.2	2.3	-	2.5	4.7
MB	2	-	-	-	-	-	-	-	1.3	2.1
TWLTL	3	-	-	-	-	-	1.8	-	1.6	3.1
CLTL	3	-	-	-	-	-	1.1	-	1.0	2.0
ALTL	3	-	-	-	-	-	1.2	-	1.2	2.6
RMD	3	-	-	-	-	-	1.4	-	1.5	3.0
MB	3	-	-	-	-	-	-	-	1.0	1.6

Table 7. Selection guide for median treatments.

	ADT					
	Low		Medium		High	
Level of Development	Median Treatment	Construction Option	Median Treatment	Construction Option	Median Treatment	Construction Option
Low	TWLTL	1	TWLTL	1	TWLTL	1
	CLTL	1	CLTL	1	CLTL	1
			ALTL	1	ALTL	1
			RMD	1	RMD	1
					TWLTL	2
Medium	TWLTL	1	TWLTL	1	TWLTL	1
	CLTL	1	CLTL	1	CLTL	1
	RMD	1	ALTL	1	ALTL	1
	ALTL	1	RMD	1	RMD	1
			TWLTL	2	TWLTL	2
			RMD	2	RMD	2
			CLTL	2	CLTL	2
			ALTL	2	ALTL	2
					TWLTL	3
					RMD	3
					MB	1
					ALTL	2
					CLTL	3
High	TWLTL	1	TWLTL	1	TWLTL	1
	CLTL	1	ALTL	1	ALTL	1
	RMD	1	CLTL	1	CLTL	1
	ALTL	1	RMD	1	RMD	1
	MB	1	TWLTL	2	TWLTL	2
			RMD	2	RMD	2
			MB	1	ALTL	2
			ALTL	2	CLTL	2
			CLTL	2	MB	1
			TWLTL	3	TWLTL	3
			RMD	3	RMD	3
			MB	2	ALTL	3
			ALTL	3	MB	2
			MB	3	CLTL	3
			CLTL	3	MB	3

treatments and construction options are warranted. Generally, median treatments that require pavement widening are warranted only for highways that have traffic volumes >5000 vehicles/d. Median treatments that require both pavement widening and right-of-way acquisition are warranted for only two types of sites: (a) highways that have traffic volumes >5000 vehicles/d and driveway densities >60 in 1.6 km and (b) highways that have traffic volumes >15 000 vehicles/d and driveway densities >30 in 1.6 km.

The two-way left-turn lane (option 1) is the most desirable median treatment in all cases considered. The median treatments decrease in benefit-cost ratio in about the order that they have been presented throughout this paper. The continuous left-turn lane is dominated by the two-way left-turn lane; i.e., in all cases where a continuous left-turn lane could be used, a two-way left-turn lane would be better. This finding results from the assumption that the continuous left-turn lane has the same effectiveness as the two-way left-turn lane but a higher cost. Therefore, the continuous left-turn lane should not be used unless there is direct evidence that it is more effective than the two-way left-turn lane.

The selection guide given in Table 7 provides an excellent basis for choosing among alternative median treatments when no better information is available. However, the user should be aware of the limitations of the guide imposed by the methods used in its development. The benefit-cost ratios are based on typical values of construction cost, accident reduction, and delay reduction. The estimates for accident and delay reduction for several median treatments are based on assumptions that appear reasonable but cannot be completely supported by research results within the current state of the art. If the user can estimate these quantities for a particular site, a more reliable evaluation will result. In this case, however, the benefit-cost procedure of this paper provides a useful framework for evaluating the available alternatives.

The user should also recognize that some important considerations are beyond the scope of an economic analysis but may well have an important impact on the final decision. For example, the economic analysis does not completely reflect the role of operational flexibility in evaluating median treatments. An arterial highway with a two-way left-turn lane is far more flexible operationally than a highway with a median barrier. Such flexibility makes routine operation less restrictive since left-turns are not prohibited, and the treatment has better service capability under transient conditions such as roadway construction or a traffic accident. In this case, both the economic and operational considerations favor the same median treatment, but in other situations there may be trade-offs to be made by the decision maker. In short, the economic analysis is an extremely important part of the selection of an optimal median treatment, but other less quantifiable factors also deserve consideration.

ACKNOWLEDGMENTS

Much of the analysis in this paper was performed by Midwest Research Institute as part of a Federal Highway Administration contract. We acknowledge this sponsorship but also state that the findings and conclusions of this paper are our own and do not necessarily represent the views of the Federal Highway Administration.

REFERENCES

1. J. A. Azzeh, B. A. Thorson, J. J. Valenta, J. C. Glennon, and C. J. Wilton. Evaluation of Techniques for the Control of Direct Access to Arterial Highways. Federal Highway Administration, U.S. Department of Transportation, Rept. FHWA-RD-76-85, Aug. 1975.
2. J. C. Glennon, J. J. Valenta, B. A. Thorson, and J. A. Azzeh. Technical Guidelines for the Control of Direct Access to Arterial Highways—Volumes 1 and 2. Federal Highway Administration, U.S. Department of Transportation, Repts. FHWA-RD-76-86 and FHWA-RD-76-87, Aug. 1975.
3. Manual on Uniform Traffic Control Devices for Streets and Highways. Federal Highway Administration, U.S. Department of Transportation, 1971.
4. T. H. Mulinazzi, and H. L. Michael. Correlation of Design Characteristics and Operational Controls With Accident Rates on Urban Arterials. Joint Highway Research Project, Purdue Univ. and Indiana State Highway Commission, 1967.
5. J. W. McDonald. Relation Between Number of Accidents and Traffic Volume at Divided Highway Intersections. HRB, Bulletin 74, 1953.
6. G. M. Webb. The Relationship Between Accidents and Traffic Volumes at Signalized Intersections. Proc., ITE, 1955.
7. An Evaluation of a Safety Project. Traffic and Safety Division, Michigan Department of State Highways, Lansing, Dec. 1968.
8. Evaluation of Four Safety Projects. Traffic and Safety Division, Michigan Department of State Highways, Lansing, TSD-G-207-2, 1972.
9. J. C. Ray. Evaluation of a Two-Way Median Left-Turn Lane. Traffic Engineering, March 1961.
10. R. B. Sawhill and D. R. Neuzil. Accidents and Operational Characteristics on Arterial Streets With Two-Way Median Left-Turn Lanes. HRB, Highway Research Record 31, 1963, pp. 20-56.
11. R. D. Lewis and others. A Study of Two-Way Left-Turn Lanes by Technical Council Committee No. 10. ITE, Technical Notes, Dec. 1976.
12. M. R. Hoffman. Two-Way, Left-Turn Lanes Work! Traffic Engineering, Aug. 1974.
13. R. C. Thomas. Continuous Left-Turn Channelization and Accidents. Traffic Engineering, Dec. 1966.
14. Need to Improve Benefit-Cost Analyses in Setting Motor Vehicle Safety Standards. National Highway Traffic Safety Administration, U.S. Department of Transportation, July 1974.
15. T. C. Thomas and G. I. Thompson. The Value of Time Saved by Trip Purpose. Stanford Research Institute, 1970.

Publication of this paper sponsored by Committee on Operational Effects of Geometrics.

Highway Design Consistency and Systematic Design Related to Highway Safety

John C. Glennon, Transportation Consulting Engineer, Overland Park, Kansas
Douglas W. Harwood, Midwest Research Institute, Kansas City, Missouri

This paper proposes a more systematic approach to highway design for achieving consistent designs to meet the needs of drivers. It is intended as a catalyst toward promoting optimal improvements of existing facilities. Its nature is conceptual. The topics covered include (a) a critique of current practices; (b) the evolution of highway design; (c) objectifying the design process; (d) consistency of design in relation to driver expectancy; (e) application to achieve design consistency; and (f) developing a cost-effectiveness methodology.

For almost 4 decades, highway designers have relied on criteria presented in a series of design policies of the

American Association of State Highway and Transportation Officials (AASHTO). Although these publications provide a unique framework for geometric design, they neither treat geometric design as a systematic process nor provide any insights on designing highways to meet the critical needs of drivers. The AASHTO design policies have often led to inconsistently designed highways. Conceived as a way of communicating standards of good practice, these policies have often become the sole authorities. When asked about the adequacy of a design, some designers say "It's consistent with the AASHTO blue book" rather than "It meets the needs of the driver."

As the recent AASHTO "3R Guide" (1) shows, the basic scenario of the highway community is rapidly changing from a massive road-building campaign to a decided attempt to optimize the traffic safety and service of existing highways. Although many design errors are "poured in concrete," this changing emphasis provides an outstanding opportunity to improve existing highways so they are more consistent with the needs of the driver. But this goal can only be accomplished if the design process is objectified to the extent that it maximizes the effectiveness of design improvements subject to funding constraints.

EVOLUTION IN HIGHWAY DESIGN

Between the inventions of the wheel and the automobile, the primary concern of road builders was "getting the road user out of the mud." Only the structural aspects of design were considered. In the 1920s, when a personal automobile became a reality for many people, there began the evolution of a highway design technology of which many remnants remain. Most of the early highway design engineers came from railroad engineering backgrounds.

As highway transportation developed in the 1930s (aided particularly by government employment-support programs), more and more paved roads were built. By the late 1930s, the number and speeds of vehicles began to multiply. With these trends came frequent traffic jams and large increases in highway fatalities.

In 1937, as a reaction to these highway transportation problems, the American Association of State Highway Officials (AASHO) organized the Special Committee on Administrative Design Policies. The purpose of this committee was the formulation of administrative policies aimed at stimulating uniform practices of good highway design that would result in maximum safety and usefulness. Between 1938 and 1944, this committee formulated the following seven policy statements: A Policy on Highway Classification, September 16, 1938; A Policy on Highway Types (Geometric), February 13, 1940; A Policy on Sight Distance for Highways, February 17, 1940; A Policy on Criteria for Marking and Signing No-Passing Zones for Two- and Three-Lane Roads, February 17, 1940; A Policy on Intersections at Grade, October 7, 1940; A Policy on Rotary Intersections, September 26, 1941; A Policy on Grade Separations for Intersecting Highways, June 19, 1944; and A Policy on Design Standards—Interstate, Primary and Secondary Systems. Many of the criteria presented in these policies still undergird current AASHTO design policy manuals. These criteria, of course, were based on the vehicle performance, highway design, and traffic operations of the 1930s. As a result, the validity of their application in current highway design technology may be questionable.

As an example of the mismatch between design standards and current highway operations, consider the example of the design and operation of passing zones. First, the design of passing sight distance (2) only indirectly considers the design of usable passing zones. The second inconsistency is that the design for passing sight distance and the striping of highways for no-passing zones are based on entirely different criteria. The current MUTCD (3) standards for no-passing zones (which indirectly set the dimensions for passing zones) are based on criteria presented in the 1940 AASHO policy (4). Unlike the current design for passing sight distance, which uses a constant 16.1-km/h (10-mph) speed difference between passing and passed vehicles for all design speeds, the sight distance for striping is based on speed differentials that range from 16.1 km/h (10 mph) at a 43.3-km/h (30-mph) design speed to 40.2 km/h (25 mph) at a design speed of 112.7 km/h (70 mph). These criteria are considerably more liberal (and more hazardous) than the design criteria (5).

Not only is the validity of current design standards in question but, more important, geometric design problems are also compounded by the lack of a systematic approach to highway design. Present methods of design are often based on solutions to old problems rather than the specific nature of the problem at hand. In addition, because of the complexity of highway design, the design tasks are generally assigned to seemingly independent teams, which ignores the basic principles of system design optimization. Although direct lines of communication may exist between task teams, the lack of defined responsibility and authority toward the total system design may prevent a solution close to the optimum.

Recently, increasing emphasis has been given to the systems engineering approach to design. This is a creative form of problem solving that emphasizes the total design or task rather than merely considering the efficiency of each component part. The primary principle in applying the systems approach to design is to maximize system performance for a given cost or to minimize cost for a given performance. This general approach, of course, is not new. What is new is that the systems approach is completely rational rather than intuitive and uses such formalized techniques as game theory, queuing theory, linear programming, dynamic programming, control theory, critical path methods, network theory, and various optimization techniques.

In the past, the complexity of the highway systems design process often forced highway administrators to decompose the process unnaturally into noninteractive tasks, ignoring many of the necessary feedback aspects of the process. Unfortunately, the highway engineering community has not had the necessary tools to consider all of the interactions, let alone objectively weigh alternative designs, in coordinating the data and performing the design.

Now that the Interstate system is nearly complete and there is a trend toward improving the safety and usefulness of the 5 968 000 km (3 700 000 miles) of existing highway network, it is past time to develop an objective design process whereby the design engineer can both design new facilities and optimize future efforts to improve the performance of the existing system.

OBJECTIFYING THE DESIGN PROCESS

To avoid some of the highway design problems of the past requires a comprehensive description of the highway design process. In other words, the total process must be completely defined from setting goals to achieving the completed design (or redesign) of a highway. The entire, conceptualized design process is shown in Figure 1 and discussed below. An appreciation of the relations and interactions shown in Figure 1 is the first step toward making each element of the idealized design process concrete rather than abstract. A major research

Figure 1. The highway design process.

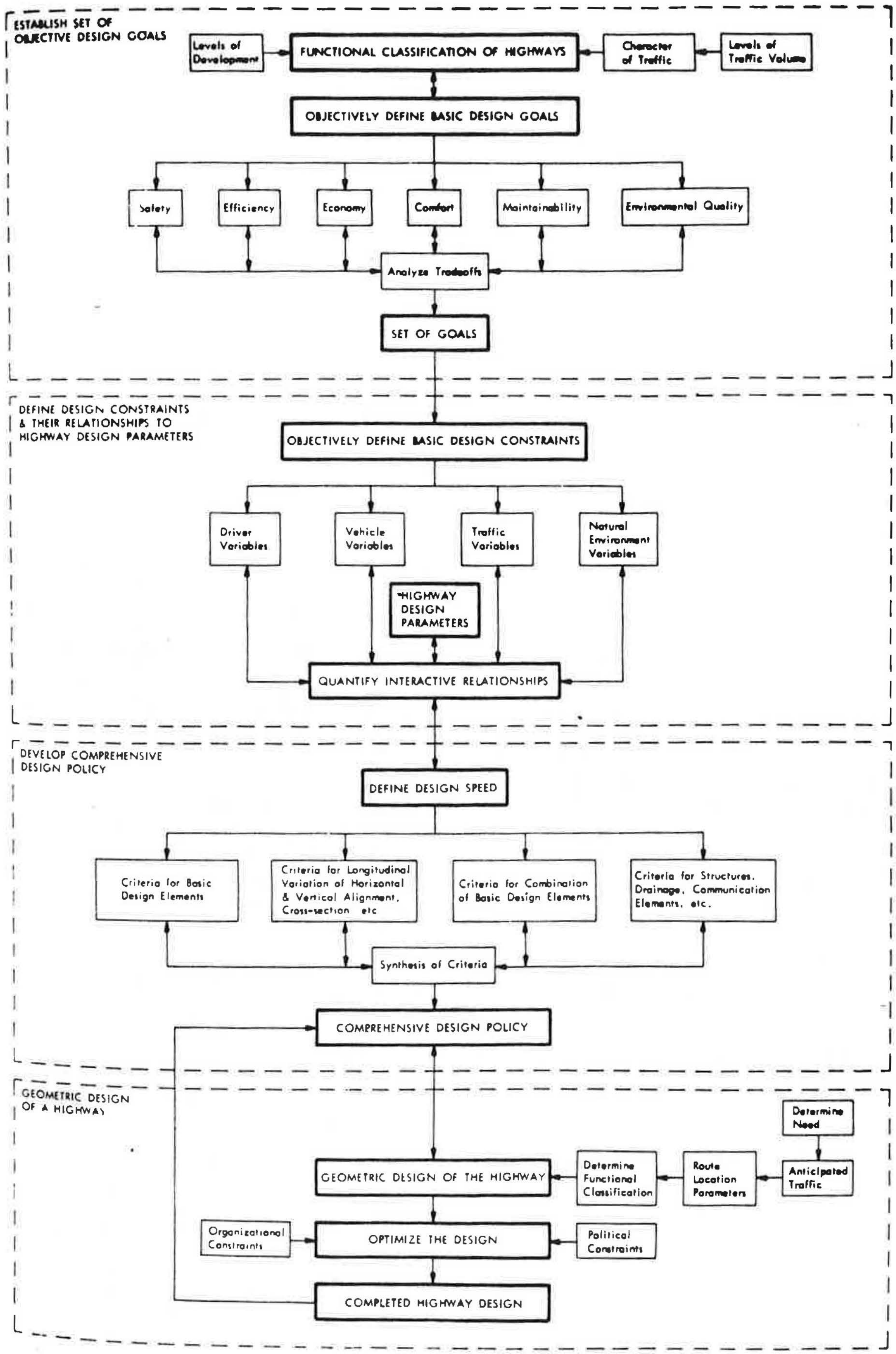
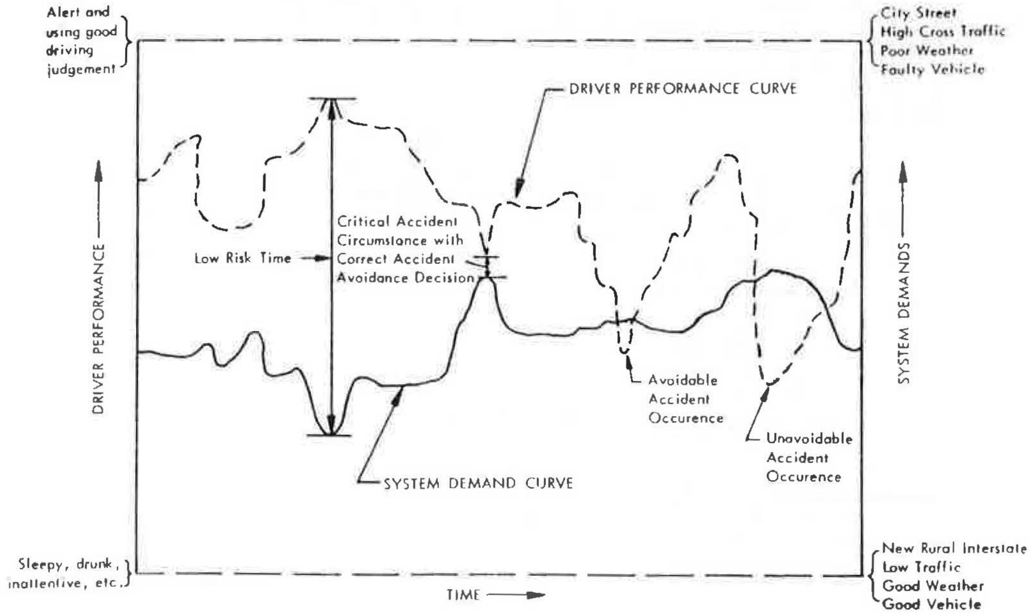


Figure 2. Conceptualized relation between driver performance and highway system demands in creation of accident circumstances.



effort is necessary to develop practical, yet valid, procedures and criteria for each element of the design process. The following discussion of the design process suggests an approach that addresses the need for specific design criteria, performance measures (measures of effectiveness), and decision-making tools.

The apex of the objective design process is the requirement that desired goals be defined and completely quantified. In addition, of course, these goals must be defined within the framework of a functional classification of highways. This points to a primary weakness of the AASHTO policies. Although they name the goals of safety, efficiency, economy, and comfort, they do not operationally define these goals.

The first part of objectifying the design process, therefore, requires a format for functional classification of highways and the formulation of a framework for operationally defining the goals of highway design in each functional class. The functional classification should consider the trade-offs between the functions of traffic service and land access, including rural or urban development needs and the level and type of traffic to be served.

The second major step in defining the design process shown in Figure 1 is an objective description of the basic constraints of driver, vehicle, traffic, and environmental characteristics and the interactive relations among these characteristics and between them and roadway characteristics. It is also important to identify how these constraints and their interactions set the requirements for the development of design criteria.

In developing design criteria that are functionally related to the design constraints, the real solution is one of matching the limited sensory and motor capabilities of the driver to the requirements of the driving task for various combinations of vehicle, roadway, traffic, and environmental constraints. Figure 2 shows conceptually that the performance of most drivers is usually adequate to the demands of the highway system. Accidents occur when either (a) driver performance falls below the level required by the system at that time or (b) system demands exceed driver performance at that time. In developing geometric design criteria, therefore, a basic principle should be to avoid peaks in the system demand curve created by inconsistency in design.

In the design process, a lack of understanding of basic design constraints and how they affect the solution contributes to piecemeal solutions that prevent optimization. The current approach tends to ignore the consistency of various combinations of design elements and thus oversimplify the process and limit the reliability of relations for most design purposes. But the primary reason for the lack of useful and definitive relations between design criteria and basic operational constraints is that these definitions depend on the complex interactions between the components of the highway transportation system, between their attributes, and between these and their environment. Until the significant interactions in the system can be quantified, reliable design criteria cannot be established.

The next major step in the design process (Figure 1) involves defining design speed as a function of design goals and constraints for each of the functional classes of highway. Without question, the "design equation" is most sensitive to vehicle speed—not only because the ability to stop or corner is a function of the square of speed but also because the impact forces of a collision are also a function of the square of speed.

AASHO design policies define design speed as "the maximum safe speed that can be maintained over a specific section of highway when conditions are so favorable that the design features of the highway govern" (2). This definition is abstract and does not lend itself to being an objective basis for design. It is difficult to imagine, under conditions "so favorable" and with modern design standards of 3.6-m (12-ft) lanes, flat cross slopes, and relatively flat grades, that any design feature other than horizontal curvature could govern maximum safe speed. Actually, in a physical sense, this is true. If driver, vehicle, traffic, and environmental constraints are eliminated from the design equation, the only design feature that physically governs maximum safe speed (for modern highway designs) is horizontal curvature. If this were true in an operational sense, the speed for long, level, tangent sections would be unrestricted and, where horizontal curvature was introduced, the concept of an overall design speed for that facility would be incongruous.

What is required is an operational definition of design speed that encompasses driver, vehicle, roadway, environmental, and traffic constraints and their relations to the design of an efficient, safe, and economical highway facility. To achieve this basis, for example, the designer requires knowledge of the characteristics of a "design vehicle" and how they relate to vehicle stability at various speeds—e.g., aerodynamics, suspension, weight, weight distribution, steer angle related to turning radius, accelerative capabilities, and braking capabilities.

The next major step is to define the design criteria objectively. The different kinds of criteria apply to the specification of the basic design elements, the longitudinal variation of horizontal, vertical, and cross-sectional elements, and the combinations of design elements (in general but also for special locations such as intersections, interchanges, and weaving sections). The process of developing design criteria involves analyzing the criticality of the interactive relations between the design constraints and the design elements for various highway speeds and selecting that level of criticality that limits the probability of an undesirable event (e.g., accident or congestion).

Synthesis is an important and necessary part of this development. Complete and comprehensive documentation of data is of little use unless it can be synthesized into a usable body of knowledge. By means of this kind of synthesis, sensitivity analysis can be performed to identify the more significant parameters that affect the safety effectiveness of any design improvement.

Figure 3 shows the general matrix of analysis. The necessary synthesis of data and information on interactive relations involves the following five basic steps:

1. Define a measure of hazard;
2. Formulate multidimensional data matrixes;
3. Statistically select an appropriate hazard value from data elements in each matrix cell;
4. Apply statistical procedures to predict expected hazard values for empty matrix cells; and
5. Reiterate the synthesis process, combining matrixes for higher orders of development.

First, we must define a measure of hazard so that the effect of varying dimensions of highway design elements, and combinations thereof, can be objectively evaluated (this step is described further below). Second, for each value of a design element, multidimensional data matrixes of hazard measures are classified by incremental values for the various combinations of the design constraints. The class range for each design element or design constraint in the matrix is then determined by analyzing the sensitivity of the dependent hazard measures to variations in the values of the design elements and design constraints. Third, within each cell of each matrix, the data elements (if there are more than one) are statistically analyzed to select the appropriate hazard value for that cell. Fourth, statistical procedures (analysis of variance, multiple regression, and so on) are applied to each data matrix to predict the expected hazard values for any empty matrix cells. And, finally, the synthesis process is reiterated, and successively higher orders of development are achieved by combining appropriate matrixes (submodels) into more inclusive matrixes.

A measure of hazard must be defined so that the effect of varying dimensions of highway design elements,

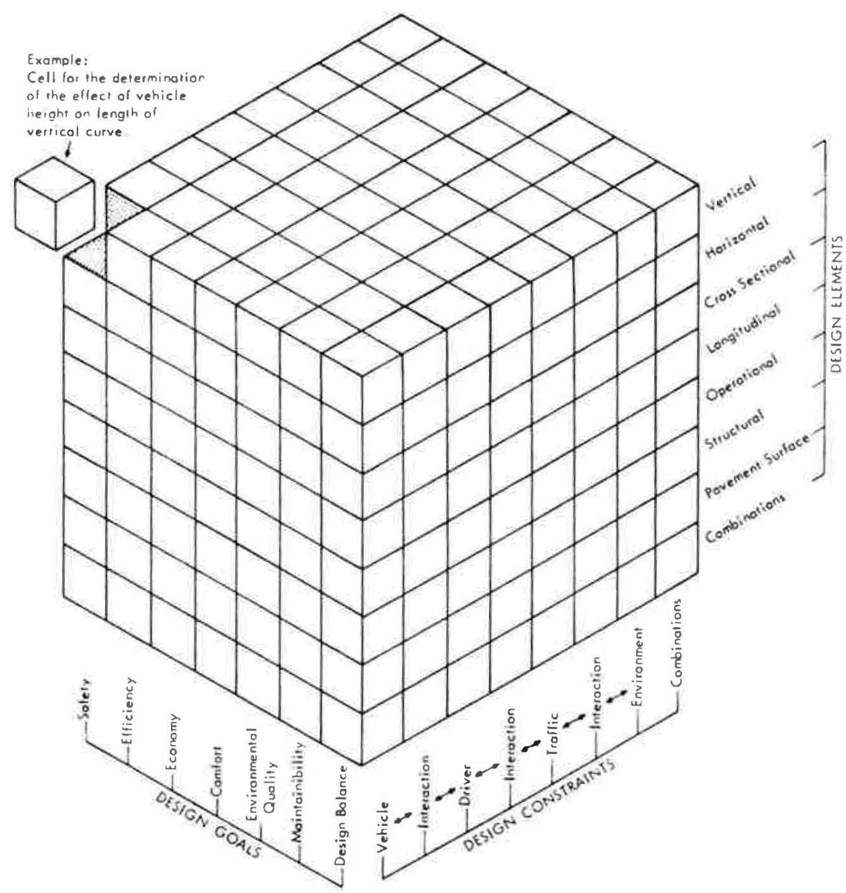


Figure 3. General matrix for synthesis of interactive relations.

and combinations thereof, can be evaluated by some criterion of "good." At any location, the degree of accident hazard is a function of two variables: accident frequency and accident severity. If two locations have the same accident frequency, the one that has the lower accident severity is less hazardous. If two locations have the same accident severity, the one that has the lower accident frequency is less hazardous. Thus, neither accident severity nor accident frequency can serve alone, but both must be integrated into one criterion.

The degree of accident hazard can be defined in several ways. It is a measure of the potential for a particular highway location to produce a given time rate of accidents with some average consequence (such as average cost or the number of fatalities, fatal accidents, or fatal plus injury accidents per total accidents). In short, the definition of accident hazard depends on the definition of accident severity, which, in turn, depends on the objective of the highway safety improvement program—whether it is intended to maximize the reduction of total accidents, accident costs, fatalities, fatal accidents, or fatal and injury accidents.

Because the process of relating all dimensional values of the design elements and the design constraints to particular values of hazard is a very complex task, it is extremely difficult to visualize the final product of the synthesis. But, for the sake of illustrating the proposed process of sensitivity analysis, let us assume that the product of the synthesis will take the form of a mathematical model that relates the independent variables that dimension the design elements, the design constraints, and the many interactions thereof. Because of this complexity, the practical application of the synthesis of interactive relations may be highly questionable. Using this kind of formulation for a practical cost-effectiveness decision-making framework may be so cumbersome as to render it useless.

The discussion above suggests that the model be tested for sensitivity to various levels of the independent variables. As the variables that contribute lesser sensitivity are discovered, they are dropped from the model, and the newer, simplified model is tested for predictive precision. This process is repeated, and the least significant variables are successively dropped or combined until the trade-off between predictive precision and simplification for practical application is optimized. The final form of the model will predict a large portion of the variation in the hazard measure by means of the simplest possible model of independent variables.

Successful development of a comprehensive set of design criteria forms the basis for an objective design policy that will enable the highway engineer to design each highway close to optimum. These designs can be accomplished if the art and the science of decision making are placed in the proper perspective, the tools of scientific decision making are brought advantageously to bear at the appropriate points in the design process, and engineering judgment is focused at the appropriate levels. In addition, the comprehensive and objective design policy will provide a framework for assimilating future improvements of design data and technology into the design process.

CONSISTENCY OF GEOMETRIC DESIGN IN RELATION TO DRIVER EXPECTANCY

Consistency has always been recognized as an underlying principle in highway design as exemplified by the following rules of thumb contained in AASHTO design policies. From A Policy on Design Standards (1945):

Sudden changes between curves of widely different radii or between long tangents and sharp curves should be avoided.

From A Policy on Geometric Design of Rural Highways (1954):

Horizontal and vertical alignment should not be designed independently. They complement each other and poorly designed combinations can spoil the good points and aggravate the deficiencies of each.

From A Policy on Geometric Design of Rural Highways (1965) (2):

The 'roller-coaster' or 'hidden-dip' type of profile should be avoided. Such profiles generally occur on relatively straight horizontal alignment where the roadway profile closely follows a rolling natural ground line. Examples of these undesirable profiles are still evident on many highways.

From A Policy on Design of Urban Highways and Arterial Streets (1973):

Curvature and grade should be in proper balance. Tangent alignment or flat curvature with steep or long grades, and excessive curvature with flat grades, are both poor design. A logical design is a compromise between the two, which offers the most in safety, capacity, ease and uniformity of operation, and pleasing appearance within the practical limits of terrain and area traversed. Wherever feasible the roadway should 'roll with' rather than 'buck' the terrain.

Although the concept of design consistency has been given substantial attention in the design policies, there is a general lack of explicit criteria for the contiguous combination of basic design elements or for the longitudinal variations of such features as horizontal alignment, vertical alignment, and cross section. Without these explicit criteria, highway designers will continue to build inconsistent geometric details into highways.

Recent attention has been focused on design consistency through the development and widespread recognition of the concept of driver expectancy. The general term expectancy relates to a stimulus-response process in which a person with an established set of ideas and concepts is presented a stimulus (visual, auditory, tactile, or other) and responds in some way to this stimulus. Although the stimulus triggers the response, the response may be either directly related or totally unrelated to the stimulus. The person's set of ideas and concepts (predisposition), which greatly influences his or her response to the stimulus, is called expectancy.

Driver expectancy relates to the readiness of the driver to respond to events, situations, or the presentation of information. If an expectancy is met, driver performance tends to be error free. When an expectancy is violated, longer response time and incorrect behavior usually result. Although driver expectancy is similar to the basic expectancy model given above, the expected situation is always changing and environmental factors are more evident, and thus the predictability of the response is reduced. That the response is to an expected situation rather than the actual situation is the vital distinction in understanding the use of driver expectancy in the design process.

APPLICATION OF DESIGN CONSISTENCY

In the most general sense, design consistency means that combination of design elements (and their dimensional specification) that does not violate the abilities of the driver to guide and control the vehicle. Therefore, the concept of driver expectancy is wholly embodied in the general definition of design consistency. In a certain sense, then, the term design consistency can almost be used interchangeably for driver expectancy.

The term driver expectancy relates a subjective appraisal of the adequacy of driver behavioral responses to particular highway situations or conditions. From this general concept is derived the idea of design consistency, which describes those combinations of geometric design elements that do not violate driver expectancies. Thus, human factors engineers, psychologists, highway engineers, and the public for that matter can generally agree that certain extreme combinations of geometric design elements constitute inconsistent design. These are the design features that usually tend to induce noticeable discomfort in the driver.

Using the concept of driver expectancy directly, however, to determine what is or is not consistent design (particularly for those design features that are close to a threshold value) presupposes that driver expectancy can be discretely quantified for a multitude of geometric design configurations. But the feasibility of this kind of quantification is questionable. There do not appear to be any studies that lend quantification (or for that matter even dimension) to the human aspect of driver expectancy. When one looks at driver expectancy as a statistical description of the driving population, a possible basis for quantification might involve observations of overt behavior such as erratic maneuvers. But there are problems in the precision and statistical description of data not to mention the complexity of an experimental design to isolate the effects of design features from the confounding effects of diverse driver, vehicle, and environmental factors. In other words, it is unclear whether it is feasible to isolate the incremental effects of design elements and features on some measure of driver expectancy in empirical studies.

An alternative is to develop criteria (from state-of-the-art syntheses) for performance elements of the driving task based on how critical they are to the safe and efficient operation of individual driver-vehicle components subject to the constraints imposed by geometric design features. In other words, establish time-distance-speed relations appropriate to maintain threshold vehicular stability (both dynamically and in an object-avoidance mode) dependent on the following limitations: driver perception, vehicle performance, driver-vehicle and vehicle-roadway interaction, and combinations of these.

In further describing this approach to evaluating design features for design consistency, it is easiest to talk about the countermeasures to driver guidance and control problems. These may be grouped into at least six general countermeasure approaches:

1. Improve driver detection—These kinds of countermeasures apply mainly to design features that do not fit patterns of driver expectancy and are also difficult to detect but cannot be improved by direct alteration. An example is to change the position of a lane drop that is just over a crest so that it is on an upgrade just downstream from a sag vertical curve.
2. Increase driver perception and response time—These kinds of countermeasures apply to design features that do not fit patterns of driver expectancy and where perception time is limited by sight obstructions. An example is to increase the distance between a crest vertical curve and a close downstream intersection.
3. Eliminate "false cue" designs—These kinds of countermeasures apply to design features that violate driver expectancy and also misguide driver control actions. A prime example is complete redesign to eliminate a side-road intersection that is tangent to a mainline curve.
4. Decrease driver guidance and control demands—These kinds of countermeasures apply to design features that violate driver expectancy in terms of perceiving the critical nature of required speed and path corrections. A typical example is providing a spiral transition to a sharp horizontal curve. Another example is increasing short taper lengths at lane drops or at lane- and shoulder-width transitions.
5. Increase driver expectancy—These kinds of countermeasures apply to design features that violate driver expectancies that are determined by immediately preceding trip experiences. An example here is building in horizontal curvature to "break up" an 8-km (5-mile) tangent section.
6. Build "relief valve" designs—These kinds of countermeasures apply when all other countermeasures are unfeasible. For example, a lane drop can be accomplished by using a painted taper and carrying the full lane width an additional 61 to 121 m (200 to 400 ft) downstream.

Developing this kind of basis requires that performance criteria answer the following kinds of questions:

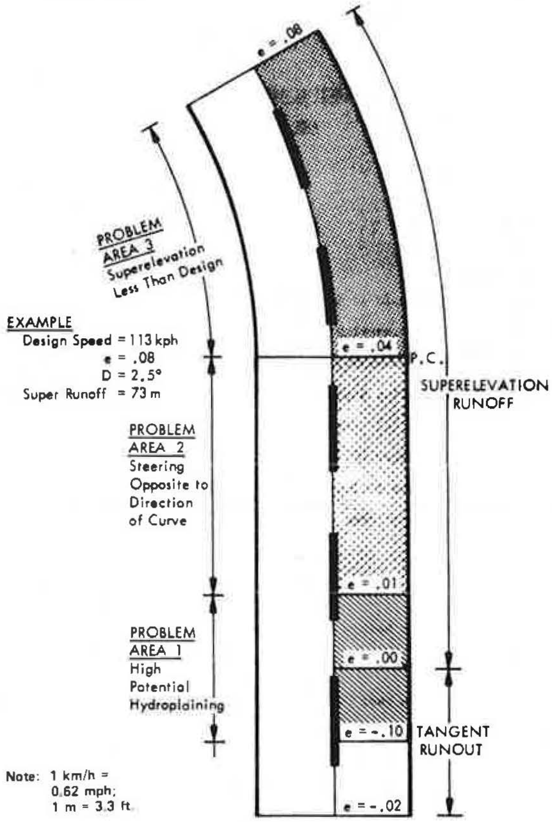
1. What are the threshold values of factors that limit perception of a geometry—e.g., lateral rate of convergence and flat line of vertical sight (parallax)? What is their relation to speed?
2. What are realistic perception times for various design features? Is perception time related to speed?
3. What maximum dynamic response (onset rate of lateral acceleration, braking deceleration, and so on) should be designed for? How does the critical nature of the responses of the driving population relate to various design features?
4. What is the time-distance degradational effect of consistent design on driver expectancy when an inconsistent design feature is introduced?

As a simple and straightforward illustration of this approach, consider the driver traversing from a tangent section to a circular horizontal curve. For horizontal curves, it is standard practice to provide a cross-slope transition from the normal crown on the tangent to full superelevation on the curve. Without a spiral transition, however, this cross-section transition appears to create a compound dilemma. This is most easily illustrated by the example shown in Figure 4. A driver approaching an unspiraled curve is presented first with problem area 1 in which the cross slope is less than 0.01 m/m (ft/ft). Because of this slight cross slope, the pavement does not drain well, and thus a section is created that has a high potential for hydroplaning. The driver no sooner gets through problem area 1 (where he or she may have experienced partial loss of control) than he or she is presented with problem area 2. In problem area 2, the driver may experience some steering difficulty because the cross slope requires steering opposite to the direction of the upcoming curve. When the vehicle passes from problem area 2 to problem area 3, the driver must reverse steering to follow the curve. At this point, if the driver steers the degree of highway curve, the lateral acceleration will be greater than that designed for since problem area 3 does not have full superelevation.

At design speed, for this example, the driver proceeds through the "compound dilemma area" in 2.6 s. Whether a driver can react adequately to these demands on his or her abilities of perception, guidance, and control in the time required is questionable.

If the example is carried one step further, past research (6) shows that drivers do not always expect vehicle stability requirements to be critical on sharper

Figure 4. Cross-slope transition area and related maneuvering problems.



horizontal curves that do not have spiral transitions. As a result, lateral accelerations on sharper curves can exceed assumed design values by as much as 2.13 m/s^2 (7.0 ft/s^2). In addition, because of insufficient space to perform an adequate spiral maneuver (the natural path of the vehicle), the rate of change of lateral acceleration can easily exceed 4.57 m/s^2 (15.0 ft/s^2).

This onset rate is clearly in the range of dynamic instability when one considers the extreme control requirements placed on the driver and the marginal ability of the tire-pavement interface to counteract such extremes.

Why, then, not add spiral transition curves that duplicate the natural path of the vehicle in a noncritical maneuver mode? Even the spiral designs that provide a continuously decreasing radius over the nominal lengths suggested by AASHO (2) will hold the onset rate of lateral acceleration under 0.91 m/s^2 (3.0 ft/s^2)—a rate that is entirely within the stable control range.

The basis of design that is apparent here is that those curves that generate more than some minimum onset rates should have spiral transitions. Although the AASHO "blue book" (2) and most state highway design manuals suggest using spiral transitions, the astute observer is hard-pressed to find a spiral curve in most states. What is apparently needed is a clearer description (sales job) for designers of the critical nature of driver control needs in the absence of the spiral.

DEVELOPING A COST-EFFECTIVENESS METHODOLOGY

Although the development of valid cost-effectiveness evaluation techniques is difficult without the kinds of inputs discussed earlier, these inputs are of little use without the development of an objective cost-effectiveness methodology for the implementation of appropriate design alternatives.

The administrator of a highway department, faced with the task of reducing accident hazard on a jurisdictional basis, must make decisions on the nature of the roadway and desired roadside designs while subject to constraints that affect those decisions. Normally, the principal constraint is limited funds. If there were no funding limitations, certainly the administrator would provide adequate lane widths, large-radius curves, long vertical curves, flat grades, and flat roadsides free of fixed objects close to the roadway. In this situation, the administrator has few decision-making problems. But, in reality, the administrator rarely works with unlimited funds and therefore strives for a strategy that allows the greatest benefits for available funds.

The basis of a cost-effectiveness analysis is that alternative methods are available for reaching an objective and each alternative requires resources and produces benefits. A cost-effectiveness analysis systematically examines the cost and effectiveness (by using some dimensional measures) of alternative methods for accomplishing an objective.

The desired cost-effectiveness methodology requires a complete decision framework for (a) computing the accident hazard associated with any highway location, dependent on the dimensions of its design elements and operational parameters; (b) computing the relative hazard reduction of alternative designs; (c) computing the total cost of a design improvement, including initial costs and differential operational and maintenance costs; (d) computing the relative cost-effectiveness of alternative designs; and (e) choosing the appropriate alternative.

The cost-effectiveness formulation has two basic components, the hazard evaluation and the cost evaluation, as shown below:

$$C/E = C_I/(H_B - H_A) \tag{1}$$

where

C/E = cost-effectiveness,
C_I = cost of improvement,
H_B = hazard before improvement, and
H_A = hazard after improvement.

As seen in this basic formulation, the hazard evaluation is used twice to compute the hazard reduction. The basic form of the hazard evaluation was discussed earlier. A brief description of the cost evaluation is given below.

In many highway situations, the difference in hazard between design alternatives may be marginal. If this is true, then at least in some cases the cost-effectiveness comparison will be most sensitive to the cost differences between design alternatives. The most important aspect of this sensitivity is the trade-off between the differentials of initial installation costs and maintenance costs. A generalized form of the cost-evaluation model is given below. Although a much more comprehensive form of the model can be anticipated, this example shows the overall concept:

$$C_A = C_I(CRF)^i_i + (C_{MA} - C_{MB}) + [C_{RA}(N_A) - C_{RB}(N_B)] \tag{2}$$

where

C_A = total net annual cost of design improvement;
C_I = total initial cost of design improvement including costs of design, right of way, removal, grading, paving, and structure and highway-user cost differentials during construction (dollars);

CRF = capital recovery factor;
t = life of design improvement (years);
i = investment return rate (percent/year);
C_{MA} = annual normal maintenance cost after improvement including surface repair and resurfacing, repainting traffic markings, mowing, snow and ice removal, and so on (dollars/year);
C_{MB} = annual normal maintenance cost before improvement (dollars/year);
C_{RA} = annual accident repair costs (to guardrails, bridges, signs, light poles) after improvement (dollars/year);
C_{RB} = annual accident repair costs before improvement (dollars/year);
N_A = annual number of collisions with highway structures after improvement: and
N_B = annual number of collisions with highway structures before improvement.

The cost-effectiveness methodology developed should, by design, lead directly to implementation. The methodology, of course, must be applied within the technical-economic-political decision-making framework of each highway agency, which ranges from the rural township highway department to the state highway department in the most urbanized state. Therefore, the methodology requires a flexible optimization strategy that is responsive to program inputs that vary according to the highway design goals of individual agencies.

The complete cost-effectiveness methodology should have several built-in decision processes other than the basic cost-effectiveness computation. The best use should be made of decision tools such as game theory, linear programming, dynamic programming, control theory, network theory, and various optimization techniques. Furthermore, the methodology should be "compartmentalized" so that subelements can be appropriately altered according to user needs without having to alter the entire methodology.

The development of the description of the design process, which was discussed earlier, will be valuable in identifying many aspects of the complete decision process. These include

1. The integration of the design goals and the highway functional classification into the decision process;
2. The ability to compare design improvements with alternative traffic operational improvements [for example, in many cases the application of "positive guidance" devices (7) may be much more cost-effective than design alternatives such as widening bridges];
3. Incorporation of decision-theory techniques to handle factors of uncertainty;
4. Methods for developing a simpler decision process by using a particular set of solutions of the cost-effectiveness methodology for a given set of input parameters;
5. The ability to optimize "earmarked" improvement programs that are based on subjective decisions or objectives other than safety;
6. The flexibility to accept future refinements in the precision of the interactive relation: and
7. The ability to balance the trade-off between precision and generalization for any particular user.

In relation to item 3—incorporating decision-theory techniques to handle factors of uncertainty—some degree of uncertainty usually affects most of the variables combined in evaluating alternative designs. Sometimes this uncertainty is dealt with by combining "conservative" values. In other cases, the best estimate value is selected for each variable. The decision-theory approach recognizes that the choice has to be made and seeks to structure the problem to incorporate estimates of uncertain factors rather than ignoring them.

In relation to item 7—balancing the trade-off between precision and generalization for any particular user—there is always a trade-off between the degree of precision and the degree of generalization in programming highway safety improvements. Maximum precision requires identifying exact values of all parameters that influence accident hazard. Implementation requires that insignificant parameters be ignored and significant parameters be categorized to minimize the collection of input data. The methodology, therefore, necessarily includes sensitivity analysis at several points in the evaluation. This analysis tests the sensitivity of the decision variable to proposed omissions and generalizations of the input parameters and provides a framework for balancing precision and generalization.

Another integral part of the proposed methodology is outlining the ways and means of implementing the total highway improvement program, as described below.

Implementing the predictive cost-effectiveness program does not mean that a spot-improvement program that uses high-accident-frequency identification procedures should be discarded. Both programs are desirable. The cost-effectiveness program identifies potentially hazardous locations; the high-accident-frequency identification program identifies locations that have demonstrated a high degree of hazard that may or may not be identified in the cost-effectiveness program. Because the cost-effectiveness program cannot precisely account for every single variable that contributes to accident hazard at every particular highway site, certain locations may actually have a higher degree of hazard than that assigned by the cost-effectiveness program. To identify these specific locations, the spot-improvement program may be more appropriate. Then, too, the cost-effectiveness methodology should be helpful in determining the best alternative improvement for sites identified in the spot-improvement program.

Unlike the spot-improvement program, which requires a comprehensive inventory of accident records, the predictive hazard approach requires a comprehensive inventory of site parameters to identify and rank potential improvement sites. Although this inventory could be the most difficult aspect of the implementation program, it may not be as difficult as it first appears. This is where the trade-off between precision and generalization comes into play. The kinds and precision of inventory items should be generalized (simplified) to a degree consistent with the desired level of program precision. It is not necessary to inventory all highways in a jurisdiction before implementing the program. A priority inventory plan can be adopted that accounts for the most sensitive variables in the hazard evaluation. In other words, the inventory plan would assign higher priorities (and hence earlier scheduling) for inventorying high-volume highways, high-speed highways, and high-hazard locations such as intersections.

To determine the general requirements for program funding, statistical procedures can be applied to obtain a representative sample inventory of hazardous locations. By using this sample to generate an estimate for the total population of cost-effective site improvements, the total program funding requirement can be estimated. This indicates to the administrator the general levels of funding that will be needed to meet various program objectives (e.g., the degree of safety payoff over specific periods of time).

SUMMARY

This paper suggests the development of a very comprehensive systems analysis for quantifying the relations between highway design elements (and their combinations) and highway safety. It also suggests the need for developing a rational cost-effectiveness methodology for optimizing the safety payoff of geometric design improvements.

The paper is critical of current AASHTO design policies and, at the same time, is "idealistic" about the potential improvement of these policies. This stance is not intended to sound pretentious but to encourage optimism toward future improvements in the design process. Only by a critical review of current practices can we ever hope to identify the missing links in achieving design consistency. On the other hand, with an idealistic attitude, we can set the highest possible goals for the future, goals that will only be modified by real (and not imaginary) constraints. The antithesis—setting shortsighted goals—would prevent the achievement of solutions that are even close to optimal.

This paper has stressed the need for more sophisticated analysis and decision-making procedures. There is little question of this need for, as the highway community strives more and more for optimality, the tools must necessarily become more objective and complex. This paper, however, does not subscribe to the "blackbox" philosophy. The methodology proposed is only a tool and as such must be comprehensible and responsive to the needs of a wide variety of users.

Future design guides must "sell" themselves to the design engineer. Traditionally, the highway design engineer has not directly accounted for the critical nature of the driver's guidance and control needs. The engineer needs to be convinced that this approach to design is not only rational but highly justified. This requires a clear and concise justification of the human-factors criteria that are used in design procedures.

The proposed methodology should be of great value in the design of new facilities as well as in the upgrading or redesign of existing highways and streets. When funds are not available for extensive upgrading of an existing facility, the methodology should aid in demonstrating the cost-effectiveness of upgrading by replacement during normal maintenance procedures. For example, for roadside hazards, as these elements wear out or are damaged or destroyed, they can be replaced by their more cost-effective counterparts.

In these times of increasing litigation against highway departments on the basis of their safety responsibility in highway accidents, a comprehensive and active program of implementing the most cost-effective improvements in order of priority should be a very convincing argument against liability for the government unit. Furthermore, this comprehensive and active implementation program should demonstrate to legislatures and the public the wise use of public funds and thus help avert the predicted deterioration of the existing system because of increasingly inadequate maintenance funds.

REFERENCES

1. Geometric Design Guide for Resurfacing, Restoration and Rehabilitation (R-R-R) of Highways and Streets. AASHTO, 1977.
2. A Policy on Geometric Design of Rural Highways. AASHO, 1965.
3. Manual on Uniform Traffic Control Devices for Streets and Highways. Federal Highway Administration, U.S. Department of Transportation, 1971.
4. A Policy on Criteria for Marking and Signing No-Passing Zones on Two and Three Lane Roads. AASHO, 1940.
5. D. W. Harwood and J. C. Glennon. Framework for Design and Operation of Passing Zones on Two-Lane Highways. TRB, Transportation Research Record 601, 1977, pp. 45-50.
6. J. C. Glennon and G. D. Weaver. Highway Curve Design for Safe Vehicle Operations. HRB, Highway Research Record 390, 1972.
7. G. J. Alexander and H. Lunenfeld. Positive Guidance in Traffic Control. Federal Highway Administration, U.S. Department of Transportation, April 1975.
8. J. C. Glennon. The Desirability of Highway Designs for Operating Speeds Above 70 mph. 40th Annual Meeting, ITE, Houston, TX, Sept. 1970.
9. D. R. Drew. Traffic Flow Theory and Control. McGraw-Hill, New York, 1968.
10. D. L. Woods. Application of Driver Expectancy in Design. Texas Transportation Institute, 1972.

Discussion

Sally Free, Center for Auto Safety, Washington, D.C.

The Center for Auto Safety agrees that the procedures now used to formulate highway design standards are not only inadequate and obsolete but also detrimental to the safety of the traveling public. Present standard-setting methodologies have failed to reduce the annual highway death toll below the staggering 40 000 mark. It is apparent that radical decreases in accident and fatality rates can no longer be realized by simply applying "common sense" solutions to old problems.

The system now in effect is one of flexible standards—that is, standards that are subject to negotiation between the Federal Highway Administration (FHWA) and the states. Vague terminology, so-called engineering judgment, and exhortatory language are poor substitutes for clear, mandatory performance criteria and objectives.

Instead of independently establishing objective performance criteria to ensure the safe design and construction of the nation's roads, FHWA has simply incorporated by reference many AASHTO design policies. One need only look at the Code of Federal Regulations, Volume 23, Part 625, to see the influence of AASHTO. The close involvement between FHWA and AASHTO during all stages of the decision-making process has had a tremendous effect on how agency decisions are made—important decisions that affect the public's safety, pocketbook, health, and environment. Indeed, this unique partnership has allowed AASHTO to shape the direction of federal standards and policies, the content of rulemaking, and even the enforcement capabilities of the regulating agency—FHWA.

This represents a rather disturbing situation since the state highway departments as a result have remained largely self-regulated, writing their own standards through AASHTO. These standards are far from optimal; much of the time the needs and safety of the driver are neglected in favor of wording that is excessively advantageous to and protective of highway officials and departments. The principal motives and objectives behind many of these design standards are the hopes of highway officials that the policies will lessen liability, decrease costs, and increase state discretionary use of federal money. These standards and policies are often simply a collection of suggestions and recommendations that are lacking in detail, are highly qualified and ambiguous, and provide so-called technical guidelines that have not always been substantiated by research or field experience. In reference to the imprecise use of terminology, one FHWA attorney has noted that "the 'standards' are so easily circumvented that they have become essentially meaningless."

The National Transportation Safety Board (NTSB) has characterized highway design standards as having originated from a "fragmented remedial approach" to safety. That is, isolated safety improvements are made based on "post mortem investigations" rather than initiating a systems approach to accident prevention at the design stage. The 1969 NTSB study, Compatibility of Standards for Drivers, Vehicles, and Highways, points to everyday traffic situations that illustrate the interrelationships of all elements in the highway system. The study maintains that the highway community has not adequately considered these interrelationships in the development and issuance of highway design standards. As a consequence, many design standards for different surface transportation subsystems are incompatible, and highway operating and design problems result. The development of performance-based design standards accompanied by a rational explanation of the function of these standards is recommended.

Although the concept of systems compatibility or systems engineering has been advanced since the 1960s, there has been little acceptance of the idea by highway departments. This lack of success stems not from a faulty methodology but rather from the need to change old policies and attitudes. First and foremost among the policy changes needed as a condition for the implementation of an effective systems engineering approach is for FHWA to establish itself firmly as a regulator and promulgate its own standards. FHWA can no longer be simply a mechanism for resource transfers. In addition, FHWA and other standard-setting agencies such as the National Highway Traffic Safety Administration (NHTSA) must better coordinate and communicate their policies, standards, and rulemaking procedures. A responsible approach to improving the process must include an extensive review and evaluation of all standards to determine their relevance and compatibility. For systems engineering to succeed, design engineers must begin thinking in holistic terms about the highway environment rather than just reacting to isolated problems that surface as a result of the present piecemeal design approach.

The systems engineering approach is advantageous in that it forces highway engineers to think in new and different terms. Isolated improvements and short-term solutions as a basic approach may be shown to be the least cost-effective alternatives over the long term. Installation expenses, maintenance costs, safety benefits, and operational efficiency can be more meaningfully evaluated when the transportation system is viewed and designed as a functional whole. An understanding of the total system operation will enable engineers to better identify and predict problems, evaluate alternatives, and implement solutions.

Systems engineering should help reduce tort liability. In recent years, the willingness of the courts to hold highway agencies and officials accountable for faulty highway design has caused tort liability to be of major concern. In an attempt to justify highway deficiencies and faulty design as accepted practice, many highway agencies are urging the adoption of lower standards. It is their hope that courts will no longer hold the highway agencies accountable to the higher standards and that this will substantially reduce their exposure to tort liability. This attempt is ill-advised and legally misguided. The cause of liability suits is not the standards but hazardous roadway conditions. A lowering of standards can only serve to increase fatalities and injuries and thereby correspondingly increase the number of claims made against highway departments. Incompatible and inadequate standards will give the lawyer the opportunity to "pick and choose" the standard that best suits the needs of a client. At a conference session on the compatibility of standards at the Fifty-sixth Annual Meeting of the Transportation Research Board, FHWA trial lawyer David Oliver reached the following conclusion:

> Without standards, accidents will occur and legal judgments will ensue. Without compatibility, standards will be not only unenforceable but also indefensible. Without cooperation there will be no 'standards.' The driver, vehicle, highway design functions must be integrated or the legal function will bare its teeth.

C. William Gray, Ohio Department of Transportation

The subject of this paper is timely, and its purpose—to promote a more systematic approach to highway design—will be enthusiastically supported by highway designers when the concept has been developed to a usable level. I have read the paper from the outlook of a designer instead of that of a researcher, since design is my background, and I believe this paper will have little impact on design until more research is performed and the system is much more thoroughly developed. The urgent need and motivation for design policy changes are clearly and accurately stated in the introduction in the statement that we are rapidly changing from a massive road-building campaign to one of improving the traffic safety and service of existing highways.

Those of us who have studied AASHTO design policies and applied them to highway design consider them to be excellent publications. If the use of AASHTO design policies has failed to meet the needs of today's drivers, perhaps the blame rests with the people who have not used these policies as a basic foundation for design and then added to that foundation from the vast store of information available from operational data and experience and current research findings. That is a very difficult task in today's rapidly changing world, and I think that is really what this paper is trying to do.

I might observe here that even our language is rapidly changing and that perhaps it does not need to change so much. Practicing highway design engineers would more readily understand and adopt new concepts and design policies if they were expressed in more commonly understood words. The last statement of the introduction is an example of how words can be hard to understand. In discussing design for the needs of the driver, it says, "But this goal can only be accomplished if the design process is objectified to the extent that it maximizes the effectiveness of design improvements subject to funding constraints." I think that says spend your money where it will do the most good. Having said that either way, we now need to go much beyond what this paper does to explain to the highway designer how to maximize design effectiveness or how to do the most good with our money.

The paper does recognize communication problems by stating the following in the summary: "The methodology . . . must be comprehensible and responsive to the needs of a wide variety of users." We could also say that the policies must be understood by highway designers so that they can apply them to all types of highway projects.

The paper recognizes, in its discussion of the evolution of highway design, the recent application to design of the systems engineering approach, which uses game theory, queuing theory, linear programming, dynamic programming, control theory, critical path methods, network theory, and various optimization techniques. Just the statement that all those things have recently evolved pinpoints the difficulty designers have in keeping current. I do not really understand some of these techniques, and I believe that many designers share my view. (Note that none of the theories named are included in the list of references at the end of the paper.)

Harwood and Glennon's discussion of objectifying the design process is based on Figure 1 but does not convey a clear understanding of the figure. At that point, it is clear that the paper will not provide a designer with an objective design process to use today, but, as the paper states, a major research effort is needed to produce a design process that will achieve safe, consistent highway designs.

The statement that "the only design feature that physically governs maximum safe speed (for modern highway designs) is horizontal curvature" should be modified to add sight distance. Maybe only horizontal curvature governs maximum speed but, in considering maximum safe speed, sight distance is a very important design feature and must be included with horizontal curvature as a governing feature.

Figure 3 shows an involved concept—a matrix for the synthesis of interactive relations—without a clear explanation.

The discussions of the consistency of geometric design in relation to driver expectancy and application of design consistency are appropriate. These subjects are of much greater concern to designers today than they were a decade ago, and they should be a major influence in future design policies.

In discussing the cost-effectiveness methodology, the authors have recognized the realistic nature of highway improvements by stating the following: "The methodology, of course, must be applied within the technical-economic-political decision-making framework of each highway agency. . . ." That has been true in the past and I am sure it will continue to be true in the future.

In view of the ever-increasing demand to improve our highways for greater traffic service and safety, it is a necessity that the highway designer have cost-effective design decision tools as proposed in this paper. I hope the paper will result in subsequent research and progress toward an early achievement of usable modern design policies, and I would encourage the authors and others to continue to work toward that objective.

Authors' Closure

We want to thank both Free and Gray for their discussions. Both of their viewpoints—Free's as a highway safety advocate and Gray's as a state highway designer—are different from our own, but their discussions help to both clarify and add depth to the intent of our paper.

Much of Free's discussion highlights the points made in our paper. She says, "The development of performance-based design standards accompanied by a rational explanation of the function of these standards is recommended," and we agree. She says, "Isolated safety improvements are made based on 'post mortem investigations' rather than initiating a systems approach to accident prevention at the design stage," and we agree. She says, "Systems engineering should help reduce tort liability," and we agree. She also says, "Installation expenses, maintenance costs, safety benefits, and operational efficiency can be more meaningfully evaluated when the transportation system is viewed and designed as a functional whole," and again we agree.

Although almost half of Free's discussion is in tune with our technical thesis, the other half gets into a far-reaching indictment of the process of setting national design standards. We definitely disagree with Free's opinion that the FHWA-AASHTO partnership has been some sort of back-room conspiracy aimed at protecting some vested interests of the state highway agencies at the expense of the motoring public. Both FHWA and AASHTO obviously share Free's deep concern for highway safety because they have (independently and jointly) sponsored many of the technological developments that have led to the improved safety performance of our highways. The roles of FHWA and the states as the major supporters of the Transportation Research Board belie Free's argument.

Free states that present standard-setting methodologies have failed to reduce the annual highway death toll below the staggering 40 000 figure. Our question is, Who ever deduced that highway design practices are the major contributor to highway accidents? Then, too, how can "present standard-setting practices" themselves ever make a measurable impact without the political recognition of the very large funding allocations needed? We must keep in perspective that it is difficult to change quickly the momentum created by hundreds of thousands of kilometers of highway that were designed and built before the advent of modern highway design technology.

What we have attempted in our paper is not to suggest discarding the present methodologies that Free claims are obsolete and detrimental but rather to recognize that the scenario of highway development has changed dramatically and that it is time to fine-tune our design methodologies so that highway agencies can reach a better balance between their safety responsibility and their fiscal responsibility. This balance cannot be achieved by using Free's "more is better" philosophy. This is why cost-effectiveness analysis is important to the design process. Although many highways could justify even high-cost safety improvements, there are other highways—particularly in the category of low-volume local roads—that cannot justify any safety improvements at all.

Gray has also highlighted many of the points in our paper but from a different perspective than Free's or ours. In addition, Gray's discussion is more a direct critique of the paper. We welcome his homespun language. He has rightly pointed to our flaws in clearly communicating our thesis. Communication is a constant problem in any profession. Idea papers often go for years without being understood by practitioners. This is mostly the fault of the authors, but then, too, the process of adapting abstract concepts so that they fit concrete and practical applications is naturally difficult and always requires considerable input by the practitioner, who usually is not paid to deal with concepts.

We also appreciate Gray's confession that he knows very little about most of the established systems engineering tools of optimization. This, of course, does not reflect on his stature as a highly respected member of the highway engineering community. Gray's statement, however, does raise a question: Why is highway design one of the few engineering professions that does not know of and regularly use these systems techniques?

In closing, we again thank the discussants for their responsive inputs. It was exactly this kind of open dialogue that we were trying to generate. Our only hope is that the dialogue will continue.

Publication of this paper sponsored by Committee on Operational Effects of Geometrics and Committee on Geometric Design.

Abridgment

Analysis of the Problem of Urban Utility-Pole Accidents

Ian S. Jones and A. Stephen Baum, Calspan Corporation, Buffalo, New York

An investigation of the problem of urban utility-pole accidents was undertaken by using 1975 data from utility-pole accidents and a sample of other urban run-off-road accidents. These data were obtained by visiting and inventorying each accident site identified in a search of police accident files in 20 urban-suburban areas included in the study.

To put the problem in perspective, the table below gives the distribution of first object struck in all single-vehicle run-off-road accidents:

First Object Struck	Number of Accidents	Percentage of Total
Utility pole	1291	21.1
Fence, guardrail	825	13.5
Sign, mailbox, parking meter, guy wire	728	11.9
Culvert, ditch, embankment	714	11.7
Tree	682	11.1
Light, signal pole	466	7.6
Fire hydrant	223	3.6
Building	215	3.5
Ground (generally rollover)	187	3.1
Wall	175	2.9
Shrubbery	120	2.0
Bridge	116	1.9
None	79	1.3
Other	303	4.9
Total	6124	100.0

Utility poles were by far the most frequent source of impact, accounting for 21.1 percent of all objects struck. Combining this figure with the fact that single-vehicle accidents accounted for 10.4 percent of all urban accidents (1) suggested that 2.2 percent of all accidents in urban areas involve impacts with utility poles.

Although it is clear that utility poles were the most frequent object struck in urban single-vehicle accidents, this is of little consequence unless the severity of such accidents relative to other fixed-object accidents is known. Distributions of injury for different objects struck in single-vehicle accidents are given below (total accidents excludes those where injury was unknown):

Object	Total Accidents	Injury Accidents Number	Injury Accidents Percent	Percentage of Total Injury Accidents
Utility pole	1166	589	50.5	31.4
Fence, guardrail	740	171	23.1	9.1
Sign, parking meter, mailbox, guy wire	668	133	19.9	7.1
Culvert, ditch, embankment	674	300	44.5	16.0
Tree	598	257	43.0	13.7
Light, signal pole	365	77	21.1	4.1
Fire hydrant	179	32	17.9	1.7
Building	163	33	21.2	1.8
Ground (generally rollover)	175	92	52.6	4.9
Wall	147	53	36.1	2.8
Shrubbery	100	7	7.0	0.4
Bridge	115	47	40.9	2.5
None	79	12	15.2	0.6
Other	202	72	35.6	3.8
Total	5371	1875	34.9	100.0

Except for vehicles striking the ground (52.6 percent), which were generally rollover accidents, utility-pole accidents had the highest percentage of injury (50.5 percent). To illustrate the overall effect of frequency and severity, this table also gives the probability of injury associated with each type of object, i.e., the likelihood of being injured by that particular object in a single-vehicle accident. It can be seen that utility poles were by far the most frequent source of injury.

The second table also shows that, in general, the proportion of injury accidents decreases as the rigidity of the object decreases. Exceptions are the categories of ground and culverts, ditches, and embankments—objects one would not necessarily associate with severe injury. However, these obstacles had a high incidence of rollover (96.3 and 20.2 percent respectively), which most likely caused the injury. Collisions that involve culverts, ditches, or embankments also had a high probability (23.6 percent) of contacting a second obstacle, which contributed to their above-average severity. The same was true for collisions with signs, mailboxes, parking meters, and guy wires; 53.8 percent of these accidents involved a second impact.

After it was established that utility poles were the most frequently struck and one of the most aggressive roadside objects, factors that differentiated utility-pole accidents from other single-vehicle accidents were examined. Few differences were noted in the variables that describe the vehicle, the driver, or environmental conditions; however, differences were detected in the variables that describe road characteristics, vehicle departure attitude, and characteristics of pole placement.

ROAD CHARACTERISTICS

It is not surprising that there was a strong cross correlation among road type, road width, speed limit, and average daily traffic (ADT). By using the combined sample of utility-pole plus run-off-road accidents, mean speed limit, mean road width, and mean ADT were calculated for each road type as given below (1 km/h = 0.62 mph and 1 m = 3.3 ft):

Item	Arterial	Collector	Local
Mean speed limit (km/h)	68.7	59.2	51.5
Mean road width (m)	9.4	9.5	7.9
Mean ADT (000s)	13.4	6.9	5.2

It is clear that road type can be characterized by using road width, speed limit, or ADT. In pursuing this further, it was also shown that ADT can be predicted from road width and speed limit so that road width and speed limit are sufficient to characterize the road system.

Figure 1. Proportion of single-vehicle accidents involving utility poles by road width and speed limit.

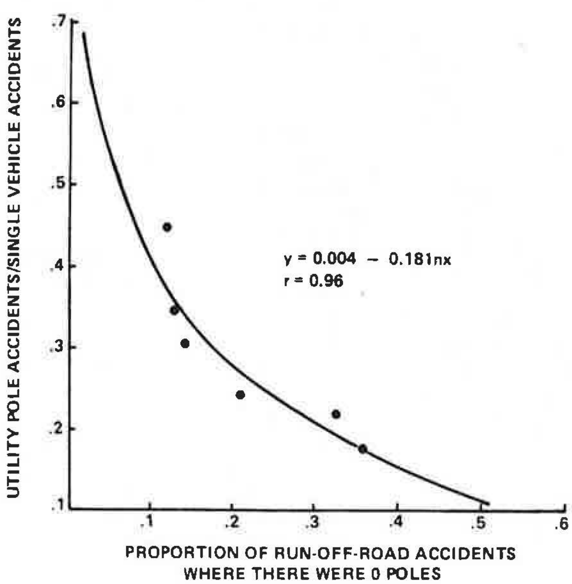
	Speed Limit (km/h)									
Road Width (m)	24	32	40	48	56	65	73	81	89	Overall
0 - 5.9	.200	.105	.105	.206	.104	.214	.120	.074	.211	.179
6.0 - 8.9	.286	.261	.208	.236	.320	.252	.168	.205	.172	.230
9.0 - 12.0	--	--	.217	.362	.491	.467	.227	.200	.143	.339
12.1 - 15.0	--	--	.175	.328	.352	.607	.318	.632	.263	.334
15.1 - 18.0	--	--	.158	.194	.333	.542	.095	.125	.030	.210
16.1 - 21.1	--	--	.286	.129	.360	.313	.219	--	.143	.214
21.2 - 24.1	--	--	--	.267	.135	.043	--	.083	.091	.122
⩾24.2	--	--	--	.500	.120	.230	.330	.077	--	.066
Overall	.236	.209	.180	.270	.280	.314	.186	.185	.146	.239

Utility pole accidents overrepresented within speed limit

Utility pole accidents overrepresented within road width

Note: 1 km = 0.62 mile; 1 m = 3.3 ft.

Figure 2. Proportion of single-vehicle accidents involving utility poles versus proportion of run-off-road accidents where there were no poles.



To show the effect of these two parameters on the frequency of utility-pole accidents, Figure 1 shows data for utility-pole accidents as a proportion of single-vehicle accidents jointly for road width and speed limit. The figures that are circled are cells in which utility-pole accidents are overrepresented compared with the overall speed-limit figure, and the figures that are boxed are cells in which utility-pole accidents are overrepresented compared with the overall road-width figure. For example, for roads that have a speed limit of 56 km/h (35 mph) and a width of 6 to 9 m (20 to 29 ft), the figure 0.320 shows that utility-pole accidents were overrepresented compared with the overall road-width figure of 0.230 and the overall speed-limit figure of 0.280. This suggests that, although there is a correlation between speed limit and road width, both variables contribute to the overrepresentation. The interaction is clear in that overrepresentation of utility poles occurs for roads with speed limits of 48 to 64 km/h (30 to 40 mph) and widths of 9 to 15 m (30 to 50 ft). This was shown to be the result of higher than average pole densities; also, roads of <9-m (<30-ft) width had high pole densities but did not have high frequencies of pole accidents, possibly because of lower travel speeds.

VEHICLE DEPARTURE ATTITUDE

The percentages of single-vehicle accidents that are utility-pole accidents are given below by travel speed (1 km/h = 0.62 mph):

Range of Travel Speed (km/h)	Utility-Pole Accidents (%)	Range of Travel Speed (km/h)	Utility-Pole Accidents (%)
0-15	13.9	64-80	15.3
16-31	16.2	81-96	22.6
32-48	23.6	97-112	26.2
49-64	24.9	113-119	41.7

The data suggest that as travel speed increases the proportion of pole accidents increases. This can be explained by a decreasing departure angle with increasing speed, which, correspondingly, increases the probability of pole contact; i.e., a vehicle exiting at a very shallow angle will have a trajectory that will expose it to more utility poles than the trajectory of a vehicle that exited at a much greater angle. A further indication of this effect is in the side-of-road-exited and road-path variables. Utility-pole accidents compared with run-off-road accidents in general had more departures to the right side of the road and a higher proportion of vehicles exiting from a straight road—situations in which one would expect a lower than average departure angle.

CHARACTERISTICS OF POLE PLACEMENT

The percentages of single-vehicle accidents that are utility-pole accidents are given below for each data collection area:

Collection Area	Utility-Pole Accidents (%)
Macon, Georgia	44.8
Knoxville, Tennessee	34.8
Columbus, Ohio	30.9
Nashville, Tennessee	24.4
Erie and Niagara counties, New York	21.9
San Diego, California	17.5

It can be seen that there is a significant variation between areas that, if one assumes that the characteristics of the driving population are approximately the same, must result from different roadway and pole-placement characteristics. Characteristics of pole placement include pole spacing, pole offset, and the number of poles within 183 m (600 ft) of either side of the struck pole or position of final rest. The latter parameter is particularly useful in that it can describe areas that have one or fewer poles.

One would expect the overall frequency of utility-pole accidents for a given area to be a function of the relative density of utility poles in that area. To test this, Figure 2 shows the proportion of utility-pole accidents in single-vehicle accidents plotted against the percentage of run-off-road accidents that occurred where there were no utility poles. Fitting a logarithmic curve through the data points shows a very strong correlation (r = 0.96) and suggests that the majority of the between-area variation is explained by the relative density of poles in each area.

Figure 3 shows the proportion of utility-pole accidents in run-off-road accidents plotted as a function of pole spacing. Fitting a regression line through the data points shows that there is a high degree of correlation (r = 0.96); i.e., as pole spacing increases, the frequency of utility-pole accidents decreases. This result complements that of Figure 2 because, from the evidence on pole spacing, sites where there were less than two utility poles had to be excluded.

Pole offset completed the definition of pole placement. Figure 4 shows the proportion of utility-pole accidents in single-vehicle accidents plotted against lateral offset at the final rest position of the pole. It can be seen that the proportion of utility-pole accidents is high at low offsets, which is where the utility poles are located. Once the mean pole offset [1.7 m (5.5 ft)] is reached, the frequency of utility-pole accidents starts to flatten out although there is still a downward trend.

REGRESSION ANALYSIS

After the factors that affect the frequency of utility-pole contact have been identified, the next step is to assess the relative importance of these parameters. This was done by using stepwise multiple regression (Table 1). At each step of the regression, the constant and coefficients of the regression equation are given together with the 95 percent confidence interval; the square of the multiple correlation coefficient is also

Figure 3. Proportion of single-vehicle accidents involving utility poles versus pole spacing.

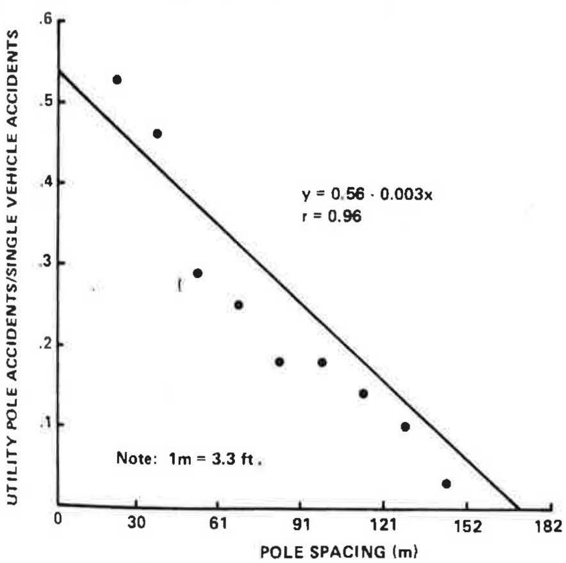


Figure 4. Proportion of single-vehicle accidents involving utility poles versus final rest position of pole.

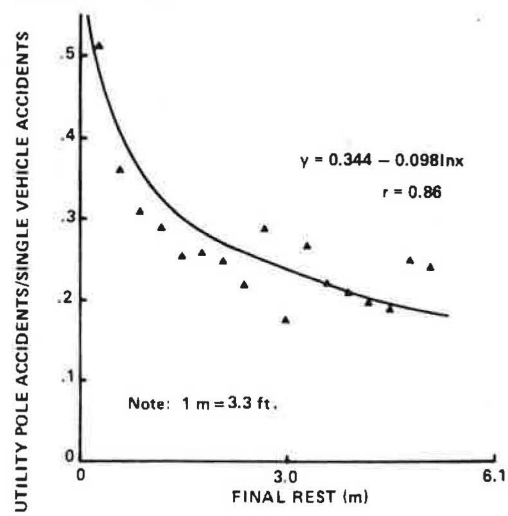


Table 1. Stepwise regression.

Step Number	Variable Entered	Constant	Number of Poles	Offset	Road Grade	Road Path	Speed Limit	Road Width	Number of Lanes	Median Width	R^2
1	Number of poles	-0.055	0.0689 ± 0.002	-	-	-	-	-	-	-	0.257
2	Offset	-0.105	0.0686 ± 0.002	0.0075 ± 0.0014	-	-	-	-	-	-	0.263
3	Road grade	-0.030	0.0682 ± 0.002	0.0093 ± 0.0014	-0.059 ± 0.013	-	-	-	-	-	0.268
4	Road path	0.0093	0.0676 ± 0.002	0.0094 ± 0.0014	-0.054 ± 0.013	-0.027 ± 0.008	-	-	-	-	0.270
5	Speed limit	0.103	0.0672 ± 0.002	0.0077 ± 0.0015	-0.053 ± 0.013	-0.026 ± 0.008	-0.0022 ± 0.0007	-	-	-	0.273
6	Road width	0.088	0.0677 ± 0.002	0.0067 ± 0.0015	-0.053 ± 0.013	-0.023 ± 0.008	-0.0026 ± 0.0007	0.001 ± 0.0004	-	-	0.274
7	Number of lanes	0.107	0.0681 ± 0.002	0.0067 ± 0.0015	-0.052 ± 0.013	-0.024 ± 0.008	-0.0025 ± 0.0007	0.0023 ± 0.0007	-0.026 ± 0.012	-	0.275
8	Median width	0.075	0.0678 ± 0.002	0.0070 ± 0.0016	-0.052 ± 0.013	-0.022 ± 0.008	-0.0021 ± 0.0007	0.0045 ± 0.001	-0.045 ± 0.013	-0.005 ± 0.0015	0.277

Note: Travel speed, ADT, pole spacing, and shoulder width deleted. 3371 data points.

given. The first variable entered is the number of poles, which explains 25.7 percent of the variation. Offset is then entered at step 2 and explains a further 0.6 percent of the variance. Road grade is entered at step 3, road path at step 4, and speed limit at step 5, and each explains an additional 0.5, 0.2, and 0.3 percent of the variance respectively. The remaining three steps given in the table each contributed another 0.1 percent to the total variation explained.

It is clear from this regression analysis that the overriding factor in predicting utility-pole accidents is the number of poles. Note that this variable not only identifies that a line of poles exists but also indicates average pole spacing since poles that were within 183 m (600 ft) of either side of the struck pole (or the rest position of the vehicle in run-off-road accidents) were counted. Furthermore, it is encouraging that offset is entered as step 2 because it complements the number-of-poles parameter by providing a more complete definition of pole placement.

The remaining parameters that are entered describe the type of road—i.e., road grade—or are related to the vehicle departure angle—i.e., road path and speed limit. This suggests that, if better measures of departure attitude were available—e.g., angle and speed—a higher proportion of variation might be explained.

REFERENCE

1. Accident Facts. National Safety Council, Chicago, 1976.

Publication of this paper sponsored by Committee on Operational Effects of Geometrics.

Abridgment

Mathematical Models That Describe Lateral Displacement Phenomena

Ali A. Selim, South Dakota State University
James L. Josey, Clemson University

In this research, a unique technique was used to collect a reliable and permanent type of data (1). Data were collected by using two super 8-mm movie cameras to study the behavior of traffic in the right lane of freeways as it approaches a vehicle parked on the right shoulder. The general tendency of vehicles as they near a parked vehicle is to swerve away from it. The path of the average vehicle at the test location is expressed by a predictive model in terms of independent variables related to geometric parameters and traffic characteristics. By using the model, the magnitude of lateral displacement at any location can be determined as the difference between the paths of the average vehicle in the presence of a side obstruction (parked vehicle) and under normal conditions (no side obstruction).

In this research, vehicles of different sizes were used and placed on the right shoulder at various distances from the freeway edge of the pavement. Vehicles were used since they are the most common type of side obstruction. A full description of the process of data collection and methods used to extract different parameters is beyond the scope of this paper but is available elsewhere (1). A brief summary of the research methodology used is presented below.

For each experiment run, a vehicle of known width was placed on the right shoulder, and the clear distance between the most remote left point of the vehicle and the edge of the pavement was measured and recorded. Two observers, each operating a camera, were signaled by a third observer by way of portable CB units to start running approximately 7.6 m (25.0 ft) of film at a speed of 8 frames/s. Three minutes of filming were designed for each experiment (1). The camera speed of 8 frames/s permitted the running of two experiments with a 15.2-m (50.0-ft) roll of film. A digital stopwatch was placed about 15 cm (6 in) in front of each camera's objective lens; these stopwatches read to $\frac{1}{100}$ of a second and appeared in the unused portion of the frame.

The first observer was stationed on a crossover (pedestrian or crossroad) and above the center of the right lane of the freeway. The observer's line of sight during filming was parallel to the traffic flow, and the edge of the pavement was ensured to be in view. The observer was completely concealed from motorists to ensure that lateral displacement did not occur because of any outside distraction but was a normal reaction of the driver when approaching the parked vehicle at the test section. A second observer, stationed evenly with the parked vehicle and on the other side of the highway, was generally outside the right-of-way; this allowed visual coverage of about 35 to 45 m (120 to 150 ft) of the roadway with the parked (test) vehicle in the middle of the observer's view.

Both films were later advanced simultaneously through stop-action projectors, and several parameters were extracted either by visual counting or by constructing special scales that were placed on the screen to measure distances. Time was read from the photographed stopwatches.

Movies taken by the first observer were used to extract parameters such as the total volume of vehicles in the right lane, including trucks and buses, and distance between the edge of the pavement and the center of a vehicle as it passed next to the parked (test) vehicle. The speeds of individual vehicles in the right lane and in the adjacent lane, headways in the right lane, and other parameters were extracted from the movie taken by the second observer.

Data from each experiment were classified as either geometric parameters (such as degree of curvature at the test location and grades in the direction of traffic flow) or traffic characteristics (such as those parameters extracted from movie films). Data were collected from two large metropolitan areas (St. Louis and Chicago) to study whether a general model could be developed that would apply to more than one met-

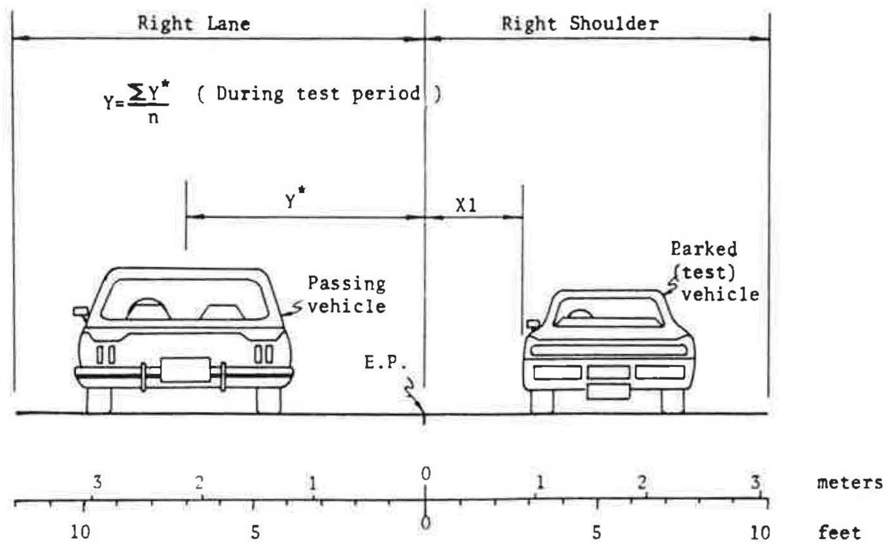


Figure 1. Path of average vehicle in relation to test vehicle.

Table 1. Numerical values of variables and their ranges in each metropolitan area.

Variable	St. Louis		Chicago		Unit of Measurement
	Maximum	Minimum	Maximum	Minimum	
Y[a]	2.18	1.64	2.17	1.85	Meters
X1	10.00	0.58	10.00	0.59	Meters
X2	1.60	-2.66	2.50	-0.50	Percent
X3	3.00	-2.00	0.00	-2.75	Degrees
X4	14.30	0.00	26.20	4.60	Percent
X5	11.10	0.00	27.30	0.00	Percent
X6	81	29	69	22	Vehicles per 3 min
X7	103.10	86.80	104.30	87.50	Kilometers per hour
X8	16	7	13	5	Vehicles per kilometer
X9	2.06	1.55	1.95	1.68	Meters
X10	0	0	1	1	Dimensionless

Note: 1 m = 3.3 ft 1 km = 0.62 mile
[a]Dependent variable

ropolitan area. Multiple regression analysis was then applied, and independently predictive models were obtained for each area as well as for the combined data to obtain a general model.

DEFINITION OF VARIABLES

One or more of the following variables appeared in the predictive models:

Y = path of the average vehicle or mean distance from the edge of the pavement to the center of the vehicle as it crosses the test location (m) (Figure 1),
X1 = distance between the most remote left point of the parked vehicle and the edge of the pavement (m) (Figure 1),
X2 = highway grade in the direction of traffic flow (percent),
X3 = degree of curvature in the direction of traffic flow,
X4 = trucks in the right lane (percent),
X5 = trucks in the adjacent lane (lane 2 in highway terminology) (percent),
X6 = volume of traffic in the right lane (vehicles /3 min),
X7 = average speed of traffic in adjacent lane (km/h),
X8 = density of traffic in the right lane (vehicles/km),
X9 = width of the test vehicle placed on the shoulder (m), and
X10 = dummy variable (only used in the general model) = 0 for St. Louis area and 1 for Chicago area.

The Y variable is dependent, and all other variables listed are independent. X1, X2, and X3 are geometric variables, and the rest of the X variables are traffic variables. Other parameters, such as the total number of lanes, lane width, and volume of traffic in the median lane (except for four-lane divided highway), appeared to be insignificant.

Table 1 gives all variables involved in the analysis and their range of occurrence in each metropolitan area.

REGRESSION AND CORRELATION ANALYSIS

Several techniques in multiple regression analysis are widely used by statisticians and engineers. The following two techniques were used because of their proven worth in the field of transportation research and especially in traffic flow analysis: (a) stepwise regression procedure (2) and (b) maximum R^2 improvement (3). The final selection of the predictive models by either technique was based solely on obtaining the best value for the multiple correlation coefficient R^2. A summary of the value of R^2 obtained by both techniques for each metropolitan area as well as for the combined data is given below (in the general model, a dummy variable is used for area identification):

Model	Stepwise Procedure	Maximum R^2 Improvement
St. Louis	0.92	0.92
Chicago	0.88	0.91
General	0.74	0.82

St. Louis Models

The stepwise regression procedure yielded model 1, mathematically described by Equation 1:

$$Y = 1.7084 - 0.0123(X5) + 0.0168(X2)^3 - 0.0256(X9/X1)^2 + 0.5736(X9*X1)^{-0.5} \quad (1)$$

Model 2 was given by the maximum R^2 improvements technique as

$$Y = 2.5031 + X5[0.0014(X5) - 0.0284] + 0.016(X2)^3 - 0.1804(X9) - 0.1023(X1)^{0.5} \quad (2)$$

This model was the best five-variable model found by the technique. The best six-variable model has a higher value for R^2. As the number of variables in the model increases, the R^2 value also increases. The best five-variable model was chosen so that the number of observations is about four times the number of variables in the model (2).

Chicago Models

The stepwise regression procedure yielded model 3, which is expressed by the following equation:

$$Y = 2.5165 - 0.6188 * 10^{-4}(X7)^2 - 0.0034(X2)^3 - 0.1216 * 10^{-3}(X5)^2 + 0.2065(X9 * X1)^{-0.5} \quad (3)$$

The maximum R^2 improvement yielded model 4 for the Chicago area:

$$Y = 2.6165 + X2[0.0737 - 0.0194(X2)^2] - 0.0453(X1)^{0.5} - 0.1083(X6/X8) \quad (4)$$

This model was the best four-variable model found by the technique. All variables in the models above were found to be significant at the 0.1 level.

Regardless of the multiple regression technique used, the models obtained for each area were found to be different in nature either in the beta coefficients or in the set of independent variables involved (X_i).

DISCUSSION OF RESULTS

In our view, the variations were mainly attributed to the following causes:

1. Unequal sample sizes were collected from each metropolitan area because of restrictions on site selection (1). Twenty sites were tested in St. Louis whereas only 16 were tested in Chicago.
2. Each metropolitan area has its own geometric and traffic characteristics that make it different from others; for example, Chicago has the following traffic and roadway features that St. Louis does not have: (a) The percentage of trucks is much higher (see Table 1); (b) there are more kilometers of depressed freeways with retaining walls; and (c) local traffic regulations, enforced by the state of Illinois, forbid trucks from using the median lanes on some sections of freeways.

An attempt was made to develop a general model by combining the data from both locations. The multiple correlation coefficient obtained by using the stepwise procedure for the combined data was 0.694. Maximum R^2 improvements resulted in a multiple correlation coefficient of 0.805 for the best eight-variable model.

The significant reduction in the multiple correlation coefficient when data were combined (compared with the R^2 value for each area separately) was expected. The reduction was mainly attributed to combining data from two different metropolitan statistical areas that are not compatible in traffic and geometric characteristics. However, an appreciable increase in the multiple correlation coefficient was obtained by using dummy variables. Dummy variables are used to account for the fact that the various areas might have separate deterministic effects on the response (dependent variable). The dummy variable (X10) had a zero value when used with St. Louis data and a value one when used with Chicago data. When dummy variables were used, the following models were obtained:

$$Y = 3.4867 + 0.0354(X2) - 0.0050(X4) - 0.178(X7)^{0.5} + 0.1491(X10) + 0.3127(X9 * X1)^{-0.5} \quad (5)$$

Model 5 is given by the stepwise procedure, and model 6 is given by the method of maximum R^2 improvements:

$$Y = 1.9932 - 0.0070(X4) - 0.0518(X1)^{0.5} - 0.3986 * 10^{-4}(X7)^2 - 0.0221(X3) + 0.0058(X2)^3 + 0.4006(X1)^{-1} - 0.0117(X9/X1)^3 + 0.1728(X10) \quad (6)$$

In using these predictive models, the average path of vehicles under normal conditions (no side obstruction) can be determined by assuming a fictitious vehicle of average width [X9 = 1.68 m (5.5 ft)] placed at a large distance from the edge of the pavement [X1 = 10.0 m (3.3 ft)].

CONCLUSIONS

The findings reached in this research are based solely on data collected from the metropolitan areas of St. Louis and Chicago:

1. From the analysis of data, it appeared that general models are not recommended for the following reasons: (a) Each metropolitan area has different characteristics related to the type of local traffic regulations, location, size, land use, social and economical status, and so on; and (b) the multiple correlation coefficients for individual area models were higher than those for the general models because the assumption that all data came from the same population holds true only for individual models.
2. In comparison with other common methods, the data collection procedure used in this study is considered one of the most economical for collecting a reliable, permanent type of data (1).
3. In our view, the maximum $\bar{R}^2$ improvement technique was advantageous over the stepwise procedure in developing predictive models.
4. The developed models can be presented graphically through a series of nomographs to show the effect of each independent variable on the amount of lateral displacement. These nomographs can provide the designer with an additional tool for analysis and comparison of proposed alternative designs.

ACKNOWLEDGMENTS

We wish to express our sincere appreciation to the individuals who devoted their time to the data collection process and analysis that made the completion of this research possible. The Civil Engineering Department of the University of Missouri at Rolla is to be especially noted for its financial support of the research.

REFERENCES

1. A. A. Selim. A Study of the Lateral Displacement Phenomena, for Uninterrupted Traffic Flow in Freeways. Univ. of Missouri at Rolla, PhD dissertation, Dec. 1976.
2. N. Draper and H. Smith. Applied Regression Analysis. Wiley, New York, 1967.
3. A. J. Barr and J. H. Goodnight. A User's Guide to the Statistical Analysis System. Department of Statistics, North Carolina State Univ., Aug. 1972.

Publication of this paper sponsored by Committee on Operational Effects of Geometrics.