TRANSPORTATION RESEARCH

RECORD

No. 1318

Safety and Human Performance

Highway Systems, Human Performance, and Safety 1991

A peer-reviewed publication of the Transportation Research Board

TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL
WASHINGTON, D.C. 1991

Transportation Research Record 1318
Price: $26.00

Subscriber Categories
II design
IV operations and safety

TRB Publications Staff
Director of Publications: Nancy A. Ackerman
Senior Editor: Naomi C. Kassabian
Associate Editor: Alison G. Tobias
Assistant Editors: Luanne Crayton, Norman Solomon
Graphics Coordinator: Diane L. Ross
Production Coordinator: Karen S. Waugh
Office Manager: Phyllis D. Barber
Production Assistant: Betty L. Hawkins

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data
National Research Council. Transportation Research Board.

Human systems, human performance, and safety, 1991.
p. cm. — (Transportation research record, ISSN 0361-1981; no. 1318)
Research papers from the 70th annual meeting of the Transportation Research Board held in Washington, D.C., in 1991.
ISBN 0-309-05158-4
1. Traffic safety. 2. Traffic signs and signals. 3. Traffic engineering. I. National Research Council (U.S.). Transportation Research Board. Meeting (70th : 1991 : Washington, D.C.) II. Series: Transportation research record ; 1318.
TE7.H5 no. 1318
[HE5614]
388 s—dc20
[363.12'5] 92-9336
CIP

Sponsorship of Transportation Research Record 1318

GROUP 3—OPERATION, SAFETY, AND MAINTENANCE OF TRANSPORTATION FACILITIES
Chairman: H. Douglas Robertson, University of North Carolina–Charlotte

Users and Vehicles Section
Chairman: Charles V. Zegeer, University of North Carolina

Committee on Vehicle User Characteristics
Chairman: Alison Smiley, Human Factors North Inc.
Secretary: Robert C. Vanstrum, 3M TCM Division
Scott L. Baker, Vivek D. Bhise, Robert E. Dewar, Mark Freedman, Peter H. Hehlen, Errol R. Hoffmann, Wolf-Dieter Kappler, P. Robert Knaff, Jan Moraal, James F. O'Hanlon, J. B. Peacock, Michael Perel, Val Pezoldt, Thomas A. Ranney, Thomas H. Rockwell, Kare Rumar, Robert R. Safford, Michael Sivak, Anthony C. Stein, Burton W. Stephens, Charles B. Stoke, E. Donald Sussman, Thomas J. Triggs, Helmut T. Zwahlen

Committee on User Information Systems
Chairman: Conrad L. Dudek, Texas A&M University System
Kenneth A. Brewer, Robert E. Dewar, David L. Helman, Gerhart F. King, David B. Knies, Harold Lunenfeld, Truman Mast, Michele Ann McMurtry, Richard A. Olsen, Jeffrey F. Paniati, Arthur W. Roberts III, Richard A. Staley, James I. Taylor, Leslie Ann Whitaker, Eugene M. Wilson, David M. Zaidel

Committee on Traffic Law Enforcement
Chairman: R. Quinn Brackett, Texas A&M University System
A. Al-Bakri, John W. Billheimer, George W. Black, Jr., Olin K. Dart, Jr., Richard L. Downey, Ted E. Dudzik, William D. Glauz, D. D. Hinton, Kent B. Joscelyn, Ruediger Lamm, Roy E. Lucke, Judson S. Matthias, Jerry G. Pigman, Ted Purtle, Thomas W. Rice, J. Michael Sheehan, Jr., Joseph J. Strojnowski, Jr., L. N. Thompson, James W. Yeasted

Committee on Traffic Records and Accident Analysis
Chairman: James O'Day, Salem, South Carolina
Secretary: C. Arthur Geurts, CAG Associates
David J. Bozak, Thomas E. Bryer, Kenneth L. Campbell, Myung-Soon Chang, Benjamin V. Chatfield, John T. Dempster, Jr., Mark Lee Edwards, Nicholas J. Garber, D. D. Hinton, Alan F. Hoskin, Paul P. Jovanis, Martin E. Lipinski, Philip P. Madonia, Joseph C. Marsh IV, Judson S. Matthias, Craig Miller, Clarence W. Mosher, Olga Pendleton, Carol Lederhaus Popkin, Rick Reischl, Robert R. Roberts, Howard S. Stein, John G. Viner, Melvyn S. Wasserman

Richard F. Pain, Transportation Research Board staff

Sponsorship is indicated by a footnote at the end of each paper. The organizational units, officers, and members are as of December 31, 1990.

Transportation Research Record 1318

Contents

Foreword vii

Part I—Speed Studies

Different Speed Limits for Cars and Trucks: Do They Affect Vehicle Speeds? 3
Herbert M. Baum, Joy R. Esterlitz, Paul Zador, and Maria Penny

Optimal Speed Limits for School Buses on Virginia's Highways 8
Jack D. Jernigan and Cheryl W. Lynn

Impact of 65-mph Speed Limit on Virginia's Rural Interstate Highways Through 1989 14
Jack D. Jernigan and Cheryl W. Lynn

Abridgment 22
Safety Impact of the 65-mph Speed Limit: A Time Series Analysis
Ronald C. Pfefer, William W. Stenzel, and Byoung Doo Lee

Use of Speed Limiters in Cars for Increased Safety and a Better Environment 34
Sverker Almqvist, Christer Hydén, and Ralf Risser

Part II—Roadway Characteristics and Accident Analysis

Analysis of Vehicle Operations on Horizontal Curves 43
Donald W. Reinfurt, Charles V. Zegeer, Brent J. Shelton, and Timothy R. Neuman

Effect of Pavement and Shoulder Condition on Highway Accidents 51
Joseph Craus, Moshe Livneh, and Ilan Ishai

Head-On, Left-Turn Accidents at Intersections with Newly Installed Traffic Signals 58
Tapan K. Datta

Influence of Road Width on Accident Rates by Traffic Volume 64
Rachel Goldstine

Identification of Accident-Prone Locations in Greater Amman 70
Fouad A. Gharaybeh

Highway Accidents: A Spatial and Temporal Analysis 75
William R. Black

Implementing a Distributed Responsibility Approach to Improving Data Quality 83
David B. Brown and Carolyn L. McCreary

Part III—IVHS Safety and Human Factors

Intelligent Vehicle-Highway System Safety: Approaches for Driver Warning and Copilot Devices 93
Anthony Hitchcock

Intelligent Vehicle-Highway System Safety: Problems of Requirement Specification and Hazard Analysis 98
Anthony Hitchcock

Engineering Design Concept for Intelligent Vehicle-Highway System Safety 104
Wei-Bin Zhang

Real-Time Responses to In-Vehicle Intelligent Vehicle-Highway System Technologies: A European Evaluation 111
Nick Ayland and James Bright

Consumer Acceptance of Adaptive Cruise Control and Collision Avoidance Systems 118
Thomas Turrentine, Daniel Sperling, and David Hungerford

Part IV—Information Systems and Driver Behavior

Factors Influencing Commuters' En Route Diversion Behavior in Response to Delay 125
Asad J. Khattak, Joseph L. Schofer, and Frank S. Koppelman
DISCUSSION, *R. D. Huchingson*, 135

Legibility and Contrast Requirements of Variable-Message Signs 137
Michèle Colomb and Roger Hubert

Distribution of Driver Spare Glance Durations 142
George T. Taoka

Foreword

The 20 papers in this Record are grouped into four general areas: (a) speed and speed limits, (b) roadway design and operations and accident analysis, (c) intelligent vehicle-highway systems (IVHSs), and (d) information systems and human performance.

Highway speed is a continuing operational and safety issue. Baum et al. examine the effect of differential speed limits for cars and trucks. Jernigan and Lynn look in Virginia at the optimal speed for school buses in one paper and the effect of the 65 mph speed limit in the next. A time series analysis was performed by Pfefer et al. to examine the impact of Illinois's 65 mph speed limit. Almqvist et al. discuss the possibilities of using a device for limiting vehicle speed.

The influence of highway design and operational practices on highway safety continues to be studied. Reinfurt et al. quantify the effects of horizontal curve features on operational variables, that is, changes in vehicle and centerline or edgeline encroachments. The effect of pavement surface, on the roadway and shoulder, on accidents in the Israeli interurban road network is reported by Craus et al. Datta analyzed crashes at intersections with newly installed traffic signals in Michigan, highlighting head-on, left-turn incidents. Does road width affect safety? The answer has considerable design and cost implications; Goldstine, after studying shoulder and road widening to 32, 36, 40, and 44 ft, compares the resulting crashes with accidents before the widening. Gharaybeh describes identification of crash-prone locations in Amman, Jordan.

Better ways to analyze crash data are always being sought. Black applies von Neumann's ratio, Moran's *I*, nearest-neighbor analysis, and a spatial-temporal autocorrelation coefficient to a transportation network situation; these are techniques not previously tried with crash data. Crash data analysis depends on data quality, and Brown and McCreary present an approach to improving data quality based on the concept of distributed responsibility.

The safety aspects of IVHS have been recognized, but the first research in this area is just coming to fruition. Hitchcock opens the discussion of IVHS safety with two papers: one looks at approaches for driver warning and copilot devices, and the other delves into problems of requirement specification and hazard analysis. Then Zhang describes an engineering design concept for IVHS safety. An evaluation of European response to in-vehicle IVHS technologies is provided by Ayland and Bright. Turrentine et al. studied U.S. consumer acceptance of adaptive cruise control and collision avoidance systems by observing focus groups.

Khattak et al. use a survey to gain understanding of factors influencing commuters' en route diversion behavior when faced with delay. In a report from France, Colomb and Hubert explain how to design variable message signs so drivers can read them more easily. Finally, spare glance distributions for one out-of-vehicle and six in-vehicle tasks are measured and modeled by Taoka.

Different Speed Limits for Cars and Trucks: Do They Affect Vehicle Speeds?

HERBERT M. BAUM, JOY R. ESTERLITZ, PAUL ZADOR, AND MARIA PENNY

During 1987 and 1988, 40 states opted to take advantage of the 1987 federal law allowing them to raise the speed limit on rural Interstates from 55 to 65 mph. The majority of these states raised the limit to 65 mph for all vehicles; however, 10 states chose a lower speed limit for trucks than for cars. Vehicle speeds were measured on rural Interstates in California and Illinois, which have a differential speed limit, and in Arizona and Iowa, which have a uniform speed limit. A posted differential speed limit on rural Interstates was found to reduce high truck speeds on the faster roads. Trucks are a smaller percentage of the high-speed traffic in states with differential speed limits than in states with uniform speed limits when average car speeds exceed 63.4 mph. Specifically, for each 1-mph increase in mean car speeds over 63.4 mph on rural Interstates, the odds relative to cars of a truck traveling about 70 mph decreases by 20 percent in the states with differential speed limits compared with states having uniform speed limits. Analysis of the mean speeds revealed that trucks travel 1.4 mph slower in states with differential speed limits than in those without. This difference increases to 3.0 mph for the fastest 5 percent of trucks.

Between 1973 and 1987, the maximum speed limit for all vehicles traveling on Interstate highways was 55 mph. In April 1987 Congress passed the Surface Transportation and Uniform Relocation Assistance Act (P.L. 100–17), which allowed states to raise speed limits from 55 to 65 mph on segments of the Interstate highway system outside densely populated areas. States may post a 65-mph speed limit on segments of the Interstate highway system that are defined by FHWA as rural or not within an area of 50,000 or more population. All other segments remain with 55-mph limits. Before 1973, many states had different speed limits based on vehicle type; for example, in Missouri cars were permitted to travel 70 mph, whereas trucks were limited to 60 mph. With the 55-mph national maximum speed limit removed, states were faced with the question of whether to set new speed limits uniformly for all vehicles or to set different limits for cars and other vehicles.

Speed limits are designed to reduce the frequency and severity of crashes by decreasing the number of vehicles traveling at excessively high speeds. The rationale for a lower maximum limit for trucks is that, from any given speed, heavier vehicles such as trucks require longer to brake and decelerate to a complete stop than do lighter vehicles such as cars (*1*). The lower limits for trucks may make stopping distances for cars and trucks more comparable and provide for a more orderly traffic flow (e.g., facilitate passing) (*2*). Opponents of the differential limit maintain that the lower limit for trucks may increase the variation in vehicle speeds and cause an increase in the number of interactions or conflicts between vehicles (*3*).

Insurance Institute for Highway Safety, 1005 North Glebe Road, Arlington, Va. 22201.

In the United States, the 55-mph national maximum speed limit was applied uniformly to all vehicles. Consequently, studies of different speed limits for cars and trucks were either done before 1973 or after 1987. A study conducted by the Civil Engineering Department at the University of Maryland (*4*) before enactment of the 55-mph national maximum speed limit found poor adherence by both cars and trucks to posted limits. At most locations with a 10-mph posted differential, the actual differential was 6 mph. On rural national routes in Ireland, where a 15-mph differential for cars and trucks was posted in 1979 (55 mph for cars and 40 mph for trucks), the observed differential was 7.4 mph (*5*).

Of the 40 states that raised the speed limit on rural Interstates, 10 set differential speed limits for cars and larger vehicles. Seven states have 65 and 55 mph limits, and the other 3 have 65 and 60 mph limits. Several states also have separate restrictions for school buses traveling on Interstates. These new driving environments allow examination of the relationship between a differential speed limit and travel speeds. A study was conducted to determine whether a differential speed limit—55 mph for trucks and 65 mph for passenger vehicles—results in a significant difference in vehicle speeds. Travel speeds were measured in states with differential speed limits and in other states with uniform speed limits, and the results were compared.

METHOD

Data

Data were collected on rural Interstates in two states with differential speed limits by vehicle type (California and Illinois) and in bordering states with uniform speed limits (Arizona and Iowa). In California the 55-mph speed limit applies to all trucks, automobiles pulling trailers, and other combination vehicles; in Illinois it applies to trucks over 4 tons, mobile homes, and vehicles pulling trailers. In both states the speed limit is 65 mph for passenger vehicles and light trucks. In Arizona and Iowa the speed limit is 65 mph for all types of vehicles.

In each state data were collected from three sites on rural Interstates for 24-hr periods in dry weather during April or May 1988. The sites were geographically and environmentally similar, and the roadway was straight and level at these sites. All sites were two-lane roads, and at least 5 mi from state borders and at least 1 mi from highway exits and entrances. Specific locations were as follows:

- *California.* Near Barstow on I-40 at 0.8 mi east of Daggett-Yermo Road (Site 1), near Indio on I-10 east of Dillion Road (Site 3), and near Boulevard on I-8 west of Route 94 (Site 4);
- *Arizona.* Near Flagstaff on I-40 at Milepost 126 (Site 5), near Yuma on I-8 at Milepost 94 (Site 4), and near Phoenix on I-10 at Milepost 94 (Site 2);
- *Illinois.* Champaign-Urbana on I-74 at Milepost 178 (Site 5), at Mt. Vernon on I-64 at Milepost 68 (Site 1), and at Effingham on I-57 at Milepost 168 (Site 3); and
- *Iowa:* Iowa City on I-80 at Milepost 222 (Site 4), near Council Bluffs on I-80 at Milepost 33 (Site 5), and near Ellsworth on I-35 at Milepost 136 (Site 3).

Data on individual vehicle speed, length, lane position, and time of day were collected using the International Road Dynamics Traffic Statistics Recorder Model 1040. When feasible, the recorder was connected to the in-pavement inductance loops maintained by the states at their permanent speed-monitoring stations. When an appropriate state monitoring site could not be located, inductive loop mats were used.

The recorder provided a raw count of the number of vehicles that crossed over at least one of the loops. In addition, the recorder provided error diagnostics, including counts of the number of passing vehicles (vehicles traveling the opposite direction to the normal traffic flow), the number of loop errors, and the number of upstream and downstream errors that occurred when a vehicle failed to pass over both mats.

A vehicle record was processed if the recorded speed was less than 999.99 km/hr, recorded within the 24-hr period, and the error code contained one of the two valid vehicle codes. The raw data were converted from metric to U.S. customary units. Records indicating a vehicle length less than 100 in. were discarded, as were records indicating a vehicle speed less than 20 mph or greater than 110 mph.

Vehicles identified from records indicating vehicle lengths of 20 ft or less, which includes cars and light trucks, were classified as cars. Records with vehicle lengths greater than 20 ft were classified as trucks, which includes buses, straight trucks, bobtails, and tractor-trailers. Occasionally, the loops had difficulty detecting trucks, possibly because of a lack of metal in the trailer or the height of the trailer from the roadway. For these sites, additional classification criteria were used. When the time lag between two successive vehicles in a lane was less than 1 sec and the first vehicle had valid speed and length measurements, the first vehicle was categorized as a truck and the second vehicle record was disregarded.

The percentages of records retained at each site are presented in Table 1. Generally, 90 percent or more of the vehicle records at a given site were used in the analysis. The one exception is the Iowa City site, for which only 59 percent of the records were used. The field engineers reported difficulty tuning the inductance loop at this site. As a consequence, the sensitivity of the loop was increased, resulting in spurious signals; thus, fewer records from the total raw count at this site were used in the analysis.

TABLE 1 PERCENTAGE OF VEHICLES WITH VALID SPEEDS AND LENGTHS BY STATE AND LOCATION

State	Location	Raw Count	Valid Vehicles	Percent
California	Barstow	4,538	4,385	97
	Indio	5,679	5,066	89
	Boulevard	3,742	3,554	95
Arizona	Phoenix	4,892	4,720	96
	Yuma	2,050	1,957	95
	Flagstaff	3,966	3,879	98
Illinois	Mt. Vernon	6,477	5,876	91
	Effingham	4,860	4,719	97
	Champaign-Urbana	9,742	9,379	96
Iowa	Ellsworth	5,871	5,491	94
	Iowa City	10,468	6,188	59
	Council Bluffs	6,107	6,002	98

Analyses

Ideally, measurements should have been made before and after the speed limit change on each of the roadways studied. However, the appropriate data from before the change were unavailable. Data gathered by the states to meet federal compliance standards do not indicate vehicle length, and the speed measurements are classified into 5-mph groups. Therefore, data from states with uniform speed limits were used as a comparison for states with differential speed limits.

In describing the data, speed distributions were compared among sites for mean speeds, selected percentile speeds, and the percentage of vehicles exceeding 70 mph. Imposing a lower speed limit on trucks was expected to reduce the number of high-speed trucks; however, it was also possible that lower truck speeds would translate into lower car speeds because of interaction among vehicles that share the same highways.

This hypothesis was tested using analysis of variance. The data were divided into five time periods: early morning (12:00 to 5:59 a.m.), morning rush hour (6:00 to 8:59 a.m.), midday (9:00 a.m. to 3:59 p.m.), evening rush hour (4:00 to 6:59 p.m.), and night (7:00 to 11:59 p.m.). The initial model treated mean truck speeds as a function of the speed limit (differential versus uniform speed limit), region of the country (West versus Midwest), day versus night, and the interaction of speed limit and region. All linear models were fitted using the general linear modeling procedure of the SAS Institute, Inc. (*6*). Weighing was not used because a small, but not statistically significant, negative relationship was detected between truck speed and number of trucks. This relationship implied that, the more trucks on the road at a given time, the lower the overall truck speeds. Therefore, weighing on the basis of sample size would have confounded the speed result.

To the extent that changes in car speeds and truck speeds among sites are parallel, any haphazard but systematic association between the 55-mph truck speed limit and average car speeds would result in biased estimates for the effect of that speed limit on truck speeds. To determine the relationship between truck speed limit and average car speeds, the truck speed analysis described previously was repeated for car speeds. Average truck speeds were plotted against average car speeds, and a positive association was noted.

Because of the association between car and truck speeds, truck speed differences among sites could be caused, in principle, not so much by lower truck speeds but to lower average speeds at a site. Hence, the corresponding car speed characteristic was used as a covariable in the analysis of truck speeds (average speed and selected percentile speeds). Separate analysis of covariance (ANCOVA) models were estimated for each speed characteristic, and each model included the differential speed limit (limit) and time of day (using the five time periods) as main effects. Car speed (speed), the covariable, was allowed to vary by whether there was a differential speed limit (speed * limit). Site variability was accounted for by using a nested effect [site(limit)].

The effect of differential speed limits on reducing the proportion of trucks among high-speed vehicles is measured using 70 mph as the speed separating low and high speeds. In order to test whether the lower speed limit was effective in reducing the proportion of high-speed trucks, an odds ratio was computed for each time of day at each site. This ratio (OR) is defined as the ratio of the odds that a vehicle traveling at or about 70 mph is a truck divided by the odds that a vehicle below that speed is a truck. This ratio can be written as follows:

$$\mathrm{OR} = \frac{\mathrm{TG/CG}}{\mathrm{TL/CL}}$$

where TG is the percentage of trucks exceeding 70 mph, TL is the percentage of trucks traveling 70 mph or less, and CG and CL are the corresponding percentages for cars. The greater the ratio OR, the greater the proportion of very fast trucks at the site.

The natural logarithm of the 60 odds ratios (four states $\times$ three sites $\times$ five time periods) were modeled using the maximum likelihood method as implemented in the CATMOD procedure of the SAS Institute, Inc. The model used is similar to that used in the ANCOVA analyses (effects included were limit, time, speed, and speed * limit). The interaction term was included because the effect of the differential speed limit laws may depend on the speed of other vehicles at the site.

The effect of differential speed limits on speed variance is not intuitively obvious. If the limits separate the traffic stream into low-speed trucks and high-speed cars, the variance would increase; however, if the limits mostly reduce high-speed trucks, the variance would remain constant or decrease. The effect of the limits on variance was estimated using an analysis of variance (ANOVA) model with day, limits, and region as main effects and the joint effect of limit by region as an interaction term.

RESULTS

The raw mean speeds by state are presented in Table 2. Trucks traveled 2.73 mph slower in the states with differential speed limits than in those with uniform speed limits. However, when adjusted for the mean speeds of cars, the difference decreases to 1.8 mph.

Using the ANOVA model, average truck speeds were estimated to be 2.7 mph less in states with differential speed limits than in states with uniform speed limits. This difference was statistically significant ($F = 12.11, p < .01$). (Estimated results refer to the least squares estimates from the model for a given speed characteristic and the specified categories of the variable.) None of the other effects (region, day, or region speed limit) were significant at the .05 level. The similar model for cars had significant effects for limit ($F = 5.13, p = .03$) and region by limit ($F = 6.01, p = .02$). The day versus night effect was not significant in either model and was dropped in subsequent analyses; however, because it is desirable to maximize the number of data points, the remaining analyses used the five time periods.

The estimated adjusted mean and percentile speeds for trucks versus observed car speeds (mean, 85th percentile, 90th percentile, and 95th percentile) from the ANCOVA models are shown in Figure 1. These estimates were slightly different from the unadjusted ones; however, the association of lower truck speeds among the states with differential speed limits remains even after accounting for differences in car speeds. The difference in estimated truck speeds was largest for the 95th-percentile speed (3.0 versus 1.4 mph for the mean speed), indicating that the differential speed limits are having the greatest effect on high-speed trucks. Speed limit was only significant for the 95th-percentile speed. Also, for each speed characteristic, the observed speeds for cars were less in the states with differential speed limits than in those with uniform limits. For example, the 90th-percentile speed for cars in the states with differential speed limits was 71.9 mph compared with 73.1 mph in the states with uniform limits.

Table 3 presents the percentage of trucks and cars exceeding 70 mph by state, site, and time period. The numerical average over all sites within a state indicates that the percentage of trucks going faster than 70 mph was twice as large in the uniform-limit states as in the differential-speed-limit states (13.8 percent in Arizona, 4.0 percent in California, 9.0 percent in Iowa, and 3.2 percent in Illinois). There was little difference by region in the percentage of cars exceeding 70 mph (26.6 percent in Arizona and 20.3 percent in California; 16.6 percent in Iowa and 18.5 percent in Illinois).

The odds ratios reflecting the likelihood of trucks exceeding 70 mph at each site differed significantly by type of speed limit, but only when mean car speeds exceeded 63.4 mph. For every 1-mph increase in speeds above 63.4 mph, the ratio of the odds ratios declines by 20 percent, implying a decrease in the proportion of trucks exceeding 70 mph in the states with differential speed limits relative to the same proportion in states with uniform speed limits. The speed of 63.4 mph represents the point at which the odds ratio in states with differential speed limits equals the odds ratio in states with uniform speed limits.

Finally, speed variance, across vehicle types, was slightly smaller in the states with differential speed limits than in the states with uniform limits (42.7 versus 44.7 mph), but the

TABLE 2 MEAN SPEEDS BY STATE AND VEHICLE TYPE

Speed Limit type	State	Trucks	Cars
Uniform	Arizona	61.1	66.4
	Iowa	62.3	65.4
Differential	California	58.2	63.8
	Illinois	59.7	65.4

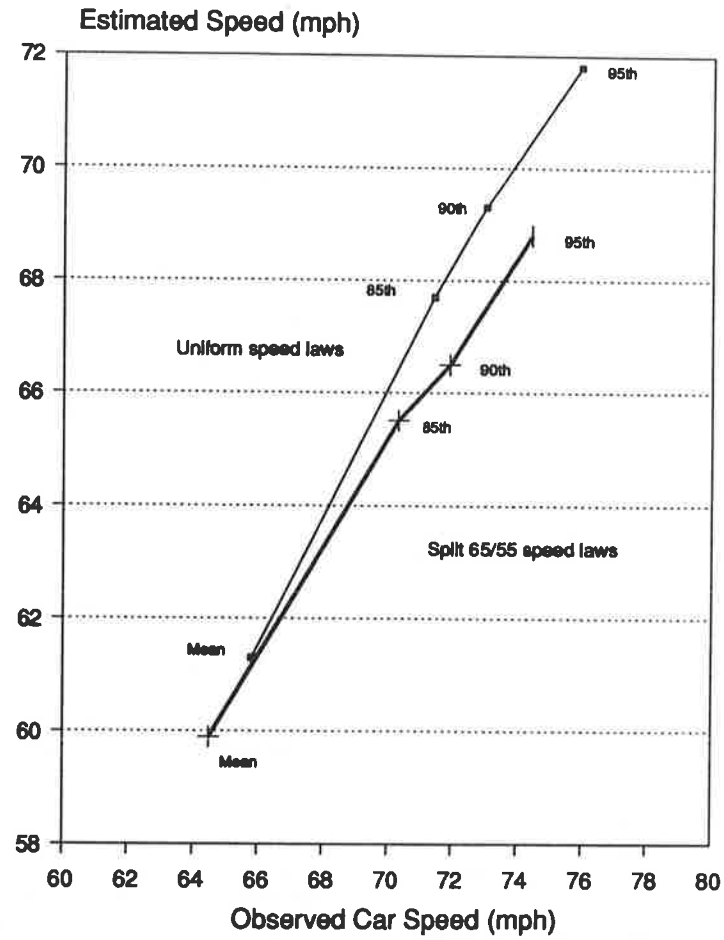


FIGURE 1 Model estimates of mean and percentile truck speeds versus comparable car speed characteristics by speed law.

difference was not statistically significant. Region of the country yields the only statistically significant effect. Speed variance in the two western states was 28 percent greater than in the midwestern states (49.1 versus 38.3).

DISCUSSION OF RESULTS

Each of the models indicates that truck speeds are lower in states with a differential speed limit for cars and trucks than in states with uniform speed limits. The magnitude of the difference varies according to which model was used. When the general traffic flow as measured by car speeds was not considered, the mean difference in truck speeds was 2.7 mph (ANOVA model), which dropped to 1.4 mph (ANCOVA model) after the adjustment was made. The former was statistically significant, whereas the latter was not. The ANCOVA model does yield a significant effect for the 95th-percentile speed. For sites with average car traffic speeds in excess of 63.4 mph, the percentage of trucks traveling very fast (70 mph or above) on rural Interstates declined dramatically in the differential speed limit states relative to the same percentage in the states with uniform speed limits.

The differences in mean speed are small; however, as expected, the main effect is in reducing the percentage of trucks exceeding 70 mph. Thus, it is likely that reducing the percentage of high-speed trucks will reduce the number and severity of highway crashes. The use of 70 mph as a cutoff is reasonable given that 70 mph is the nominal design speed for rural Interstate highways. Also, the average car speeds at all the sites (except one) for each time period (except one) were faster than the 63.4-mph threshold predicted by the model. It is also likely that the average speed will continue to increase as drivers become accustomed to the new higher limits, thereby enhancing the effect a differential speed limit could theoretically produce.

Differential speed limits may increase the variance in speeds of vehicles. However, the data do not indicate that speed variance is different in states with differential speed limit laws versus those without these laws.

Because of the large variation by site, region was discounted in importance early in the study; however, for clarity some of the final models were run separately by region. The magnitude of the effect of the differential speed limit does seem to be stronger in the Midwest than in the West, but the difference is not statistically significant and is primarily caused

TABLE 3 PERCENTAGE OF TRUCKS AND CARS EXCEEDING 70 mph BY STATE, SITE, AND TIME

			Percent ≥ 70 mph	
State	Site	Time	Trucks	Cars
Arizona	A	0:00 - 5:59	25.4	42.6
		6:00 - 8:59	31.7	40.4
		9:00 - 15:59	25.7	32.7
		16:00 - 18:59	19.6	33.4
		19:00 - 23:59	21.4	38.5
	B	0:00 - 5:59	23.1	22.7
		6:00 - 8:59	17.4	22.5
		9:00 - 15:59	9.8	23.3
		16:00 - 18:59	14.8	23.1
		19:00 - 23:59	8.7	20.4
	C	0:00 - 5:59	1.2	18.1
		6:00 - 8:59	0.0	14.5
		9:00 - 15:59	0.4	17.0
		16:00 - 18:59	0.4	24.0
		19:00 - 23:59	0.8	16.8
California	A	0:00 - 5:59	9.7	24.8
		6:00 - 8:59	11.8	33.0
		9:00 - 15:59	6.1	28.4
		16:00 - 18:59	12.2	36.4
		19:00 - 23:59	3.4	24.4
	B	0:00 - 5:59	0.7	4.0
		6:00 - 8:59	0.0	1.1
		9:00 - 15:59	0.1	0.7
		16:00 - 18:59	0.0	2.0
		19:00 - 23:59	0.2	2.3
	C	0:00 - 5:59	29.0	27.4
		6:00 - 8:59	15.4	21.9
		9:00 - 15:59	7.0	32.7
		16:00 - 18:59	5.5	30.5
		19:00 - 23:59	7.4	19.2
Iowa	A	0:00 - 5:59	2.4	10.3
		6:00 - 8:59	2.5	11.8
		9:00 - 15:59	4.7	14.8
		16:00 - 18:59	4.4	12.5
		19:00 - 23:59	9.8	14.2
	B	0:00 - 5:59	0.0	17.1
		6:00 - 8:59	15.3	36.0
		9:00 - 15:59	12.5	17.8
		16:00 - 18:59	29.4	22.3
		19:00 - 23:59	23.2	22.3
	C	0:00 - 5:59	5.8	14.0
		6:00 - 8:59	3.0	15.5
		9:00 - 15:59	1.0	13.9
		16:00 - 18:59	2.2	14.3
		19:00 - 23:59	4.8	18.1
Illinois	A	0:00 - 5:59	3.9	28.7
		6:00 - 8:59	1.9	20.7
		9:00 - 15:59	2.9	22.1
		16:00 - 18:59	4.5	26.5
		19:00 - 23:59	1.0	23.2
	B	0:00 - 5:59	4.7	18.2
		6:00 - 8:59	3.5	16.1
		9:00 - 15:59	4.4	18.7
		16:00 - 18:59	4.8	14.5
		19:00 - 23:59	5.2	19.5
	C	0:00 - 5:59	2.5	12.3
		6:00 - 8:59	3.6	20.1
		9:00 - 15:59	2.2	14.2
		16:00 - 18:59	2.1	18.2
		19:00 - 23:59	1.7	13.0

by one slow site in California. If this site is treated as an outlier, there is no evidence of a regional effect.

Lower speed limits have been found to be associated with a reduced number of motor vehicle injuries, but the relationship between differential speed limits and crashes or injuries has yet to be determined. This study has shown that differential speed limits do reduce the speeds of the fastest trucks with no apparent increase in speed variance. The next step is to determine whether this effect translates into a different pattern of motor vehicle injuries than that in states with uniform 65-mph limits.

ACKNOWLEDGMENTS

This work was supported by the Insurance Institute for Highway Safety. The authors would like to thank Scott Davis and David Harkey of Scientex for their important role in the data collection and Adrian Lund for his insightful comments.

REFERENCES

1. I. S. Jones. Truck Air Brakes—Current Standards and Performance. Presented at 28th Annual Conference of the American Association for Automotive Medicine, Denver, Colo., 1984.
2. W. S. Ferguson. *Truck Speeds in Virginia*. Virginia Highway Research Council, Charlottesville, 1968.
3. J. W. Hall and L. F. Dickinson. Truck Speeds and Accidents on Interstate Highways. Presented at 53rd Annual Meeting of the Highway Research Board, Washington, D.C., 1974.
4. *An Operational Evaluation of Truck Speeds on Interstate Highways*. Civil Engineering Department, University of Maryland, College Park, 1974.
5. R. Hearne. Car and Truck Speeds Related to Road Traffic Accidents on the Irish National Road System. *Proc., International Symposium on the Effects of Speed Limits on Traffic Accidents and Fuel Consumption*, Dublin, Ireland, 1981.
6. *SAS Users Guide: Statistics*, Version 5 ed. SAS Institute, Inc., Cary, N.C., 1985.

Publication of this paper sponsored by Committee on Traffic Law Enforcement.

Optimal Speed Limits for School Buses on Virginia's Highways

JACK D. JERNIGAN AND CHERYL W. LYNN

On Virginia's rural Interstate highways, there is a three-tiered speed limit: 45 mph for school buses, 55 mph for trucks, and 65 mph for other vehicles. On urban Interstate highways, school buses are restricted to 45 mph, but other traffic has a 55-mph speed limit. Speed theory suggests that restricting school buses to slower speeds will limit the potential severity of accidents that occur but that slower speeds may increase the probability that a school bus will become involved in a crash with a faster-moving vehicle. Forty-one states allow school buses to travel at least 55 mph on the Interstate highway system; 22 states allow school buses to travel 65 mph on rural Interstate highways. Surveys of school administrators, school bus drivers, and other interested groups indicated that a majority favor raising the speed limits for school buses to 55 mph on the rural Interstate highway system but retaining the 45-mph maximum limit on urban Interstate highways and on other systems. In 4 years there were only 17 school bus crashes on Virginia's Interstate highways, which resulted in only six injuries and no fatalities. These crashes were not attributable to the difference in speed limits or to collisions between heavy trucks and school buses. In addition, because Virginia's school buses are equipped with a speed governor that limits the maximum speed of the bus, a higher speed limit would require raising the speed allowed by the governor, which could have a deleterious effect on school bus safety on the primary and secondary systems. Thus, it was concluded that there are no compelling reasons for Virginia to raise the maximum speed limits for school buses from 45 mph and that there are reasons that caution against raising the speed limit.

On July 1, 1988, the maximum speed limit for passenger vehicles on Virginia's rural Interstate highway system was raised to 65 mph. One year later the speed limit for commercial buses (except those used as school buses) was also raised to 65 mph. The speed limits for trucks and school buses on rural Interstate highways remained unchanged, thereby resulting in a three-tiered speed limit for Virginia's rural Interstate highways: 45 mph for school buses, 55 mph for trucks, and 65 mph for other vehicles.

These changes generated some concern within the pupil transportation community that the new speed limit for passenger vehicles and commercial buses might place school buses at increased risk for accidents. Theory and traffic engineering research suggest that (a) the absolute speed at which a vehicle travels is directly related to the severity of an accident involving the vehicle and (b) the variance and distribution of the speeds of vehicles traveling on a given roadway are related to the likelihood of an accident occurring.

These relationships between speed and accidents create the presumption that the adoption of the higher speed limit could have increased the likelihood of crashes occurring between school buses and other vehicles. Officials within the Virginia Department of Education were sufficiently concerned about this possibility that they requested that a study of the effect of speed limit changes on school bus accident potential be conducted.

Although there are many possible criteria by which to assess the impact of changes in speed limits—such as convenience, economic benefits and costs associated with the resulting savings in time, and public opinion—the overriding consideration was that the optimal level of safety for students traveling in school buses be ensured. Thus, safety was the primary criterion used to assess whether school bus speed limits should be changed.

Because the safety record of school buses on Interstate highways is extraordinarily good, a conservative approach to investigating the speed limit question was adopted. Thus, compelling reasons for change would have to be indicated before any such recommendation would be made.

METHODOLOGY

An initial step in assessing Virginia's multitiered speed limits was to determine whether other states had established speed limits for school buses that were lower than those for other vehicles. Although 17 states regulate speed limits for school buses by statute, many others have provisions that allow restrictions on speed limits for school buses to be established by administrative regulation initiated by the agency responsible for overseeing student transportation in the state. It was decided that the best way to obtain comprehensive information about administrative rules was to survey the various agencies responsible for overseeing student transportation.

There is a consensus in the literature that stopping distances and crash severity are directly related to increases in travel speeds. Crash probability, however, is generally considered to be related to speed variance and the distribution of travel speeds on a given roadway. In the discussion of speed and speed variance, an analysis of the literature on crash probability and crash severity is provided, along with a description of how these issues relate to the question of establishing speed limit differentials.

In an attempt to measure the level of compliance by school buses with the 45-mph maximum speed limit, a speed survey of school buses traveling on Virginia's Interstate highways was conducted during morning and afternoon hours in the spring of 1989—hours when school buses were likely to be transporting students to and from school or special activities

Virginia Transportation Research Council, Box 3817 University Station, Charlottesville, Va. 22903-0817.

such as sporting events. A vehicle equipped with a properly calibrated speedometer followed school buses that were traveling on Interstate highways in various regions of the commonwealth. The sections of Interstate highways to be surveyed were chosen after conversations with Virginia Department of State Police division commanders, who identified general areas where school buses traveled on Interstate highways in their division's jurisdiction. The survey vehicle, which was positioned near the onramp to the Interstate highway, paced the school bus after both vehicles were up to speed. While the school bus was paced, characteristics of the bus—such as its size, whether it was loaded, and its school division—were noted. The speed traveled by the bus was measured many times between on- and offramps, and the average of these measurements was used to estimate the speed of the bus.

A recent report on school bus safety (*1*) included an extensive analysis of fatal crashes in which school buses were involved during 1982–1986. Detailed descriptions of each crash in which there was an in-bus fatality were used to analyze how the travel speeds of school buses may have been related to the crashes or their severity. Likewise, data on all crashes in Virginia involving school buses for the academic years 1985–1986 through 1987–1988 were analyzed to determine whether travel speeds of school buses were related to the crashes or their severity.

Finally, three groups (pupil transportation administrators, school bus drivers, and police agencies and other special interest groups) were surveyed. Each group was queried concerning school bus speed limits. The groups were asked several questions in common, along with a number of questions specific to their own interests and expertise.

ANALYSIS

Laws and Policies of the 50 States

Twenty-two states allow school buses to travel 65 mph on rural Interstate highways, and one state allows them to travel 60 mph. Eighteen states have established a 55-mph maximum speed limit for school buses, which represents a speed limit differential in 12 of these states. The remaining nine states have a speed limit differential for school buses. Four states allow school buses to travel 50 mph, four states allow them to travel 45 mph, and one state allows them to travel 35 mph.

In all, 28 states treat school buses the same as any other type of vehicle, and 22 have special provisions that establish a speed limit differential for school buses. Virginia is among the five states with maximum speed limits of 45 mph or less for school buses. Furthermore, Virginia and North Carolina are the only states that limit school bus speeds to 20 mph below that of most other traffic on rural Interstate highways.

Speed and Speed Variance

Many of the arguments that were made for and against differential speed limits for trucks may also apply to school buses. There are special considerations for school buses, however. Handling characteristics and occupant protection standards are different for school buses, thus affecting their crash and injury potential. These vehicles also carry students, whose lives have been entrusted to the schools. Another limitation of the comparison is that, in Virginia, the speed limit differential for school buses is sometimes 20 mph, not 10 mph as it is for trucks. Despite these differences, the truck differential analogy is useful in considering the school bus speed differential.

In the 1950s and 1960s, the use of differential speed limits was widespread in the United States, especially along the East Coast. At one time, more than half the eastern states had a differential speed limit for trucks. Virginia also has a history of establishing a differential maximum speed limit for passenger vehicles and for larger and heavier vehicles. In 1938, Virginia's first truck differential limit was imposed. The speed limit was set at 55 mph for passenger vehicles, 45 mph for trucks, and 35 mph for school buses.

The rationale on which the use of a speed differential is often based is that, if the speed for a class of vehicle is lowered, such vehicles are less likely to be involved in accidents, thereby resulting in improved traffic safety. The assumption is that the class of vehicle assigned the lower speed limit is associated with a lower level of safety when traveling at the speed assigned to other traffic. It remains to be determined whether this assumption is valid.

Although this assumption is based on the further assumption that differential speed limits reduce accident probability, previous research and accident experience have shown that this assumption is not valid. The speed that a vehicle travels is not correlated with the likelihood that it will be involved in an accident. Rather, speed is related to the severity of the accidents that occur. As the speed a vehicle travels increases, the severity of any crash involving the vehicle also increases, especially at speeds in excess of 60 mph (*2*). This assumption makes sense because the higher the speed traveled, the higher the energy that must be absorbed by the occupants and the vehicle in a collision. In fact, a 20 percent increase in speed from 50 to 60 mph results in a 44 percent increase in the kinetic energy that must be absorbed, thus dramatically increasing the severity of the consequences of an accident (*3*).

One factor that seems to affect the probability of a crash occurring is the speed of the vehicle in relation to the speed of all other vehicles on the road at the time. The greater the discrepancy between a vehicle's speed and the average speed of other vehicles on the same section of road, the more likely the vehicle is to be involved in an accident (*2*, *4–7*). When vehicles travel at widely varying speeds, the number of interactions, such as overtaking and passing, is maximized (*8*). Also, the closer a vehicle travels to the average speed, the fewer the interactions; therefore, the opportunities for a crash to occur are minimized. Thus, accident involvement rates have been shown to vary directly with speed variance, that is, how vehicles' speeds differ from the average speed. Also, the fatality rate tends to be highest for vehicles traveling at speeds that are either much higher or much lower than the average speed (*2*).

These speed characteristics are important to remember when considering the potential effects of a speed limit differential, which tends to increase speed variation. On Interstate highways, increasing speed variation would theoretically increase the number of rear-end and lane-change interactions between school buses and other traffic, thereby increasing the potential for these types of accidents.

A special consideration of the speed limit differential in Virginia relates not to travel speeds on Interstate highways, but rather to travel speeds on primary and secondary systems. Virginia's public school buses currently have a governor that mechanically limits the maximum speed of a bus to 45 mph. If the maximum speed limit were increased, the governor would have to be adjusted to allow a bus to travel at the higher speed. A governor set at a higher speed would not only permit the operation of the vehicle at that speed on Interstate highways but would also give drivers the option of traveling faster on other roadways where such speeds are both illegal and inappropriate. Thus, an advantage of the current system is that a governor now limits school bus speeds to 45 mph.

School Bus Speed Survey

Although the speed governors installed on school buses should limit school bus speeds to 45 mph or less, a speed survey was conducted to determine the actual travel speeds of school buses on Virginia's Interstate highways. Table 1 shows that the average speed traveled by public school buses in Virginia was 48 mph. Although the number of observations was small, the speeds observed for Virginia's public school buses were substantially lower than those for Virginia's private school buses and those for school buses from other states that were traveling in Virginia. The minimum speed measured for Virginia public school buses was 43 mph, and the maximum was 58 mph. Of the 42 Virginia public school buses observed, however, 9 were traveling in excess of 50 mph. Because 20 percent of the school buses were exceeding 50 mph, it is clear that at least some speed governors were not working as intended.

The three Virginia private school buses that were observed were all paced at 55 mph, as were buses from North Carolina and Maryland. A school bus from Kentucky, however, was paced at 64 mph on a rural portion of Virginia's Interstate highways on which a 55-mph speed limit for all buses was clearly posted. In fact, in Kentucky the maximum speed limit for school buses is 55 mph, so the Kentucky bus was also clearly in violation of the limit established in its home state. Although the Maryland school bus was exceeding Virginia's maximum speed limit for school buses, 55 mph is the speed that Maryland allows its school buses to travel. Likewise, although the North Carolina school bus was a yellow bus, it was clearly marked as an activity bus. North Carolina allows such buses to travel as fast as 55 mph.

One conclusion that can be drawn from these data is that appropriately geared school buses that do not have a governor or whose governor is set high enough can travel at least 55 mph on Interstate highways. It is hypothesized that a governor functioned, at least in part, to increase compliance with Virginia's 45-mph maximum speed limit for school buses. If speed limits are raised on the Interstate highways, the maximum speed allowed by a governor will also have to be raised, thus allowing buses to travel, albeit illegally, at higher speeds on all roadways.

Accident Experience

Nationally, between 1982 and 1986 there were 26 crashes involving school buses in which there were in-bus fatalities (*1*). Of these, 5 occurred on Interstate highways. Two of these crashes are of the types potentially resulting from increases in speed or from increases in speed variance. In one case, a bus traveling at an excessive speed, much higher than the prevailing speed limit at the time, crashed into a fixed object. Had the crash occurred at a lower speed, the impact velocity would have been lower and the injuries might not have been so severe. However, this bus was traveling at a speed of 75 mph, far outside the speeds being considered in this study. In another crash, a bus traveling slower than the free-flowing traffic (40 mph) was struck from behind by a tractor-trailer. This type of crash is one that may occur when the travel speeds of various vehicle types vary significantly.

Over the 3-year period before the change in the rural Interstate highway speed limit for passenger vehicles, only 10 crashes occurred on Virginia's Interstate highways. These crashes accounted for only 0.4 percent of all school bus crashes in Virginia. Thus, it is difficult to draw any conclusions from these data other than that there is not a substantial school bus crash problem for Virginia's Interstate highways. There were no fatalities reported for school buses on Interstate highways in Virginia, and Table 2 indicates that nonfatal crashes were more likely to occur when a bus was traveling at speeds of 25 mph or less, even on Interstate highways.

Table 2 also indicates that it is unlikely that there are substantial problems on non-Interstate roads related to maximum speed limit policies for school buses. On the primary highways, 83.6 percent of school bus crashes occurred at speeds of 25 mph or less, and only 4.7 percent (14 crashes) of the 3-year total occurred at speeds in excess of 35 mph. If the maximum speed limit created a problem on primary highways, it should have been manifested by an increased number of crashes occurring when the school bus was traveling at its maximum speed. Thus, a substantial speed limit–related crash

TABLE 1 SURVEY OF TRAVEL SPEEDS FOR SCHOOL BUSES ON VIRGINIA'S INTERSTATE HIGHWAYS

School Bus Type	No. Observations	Average Speed
Kentucky	1	64
Maryland	1	55
North Carolina	1	55
Virginia Private	3	55
Virginia Public	42	48

TABLE 2 SPEED OF SCHOOL BUSES INVOLVED IN CRASHES BY LOCATION

Location	Speed			
	0-25 mph (%)	26-35 mph (%)	36-45 mph (%)	46 mph and over (%)
Interstate Highway	5 (50.0)	1 (10.0)	4 (40.0)	0
Primary Highway	249 (83.6)	35 (11.7)	6 (2.0)	8(2.7)
Secondary Road	845 (85.3)	106 (10.7)	4 (0.4)	36(3.6)
City/Town Street	946 (93.5)	42 (4.2)	1 (0.1)	23(2.3)
School Facility	121 (98.3)	— —	2 (1.6)	

problem cannot be documented for primary highways from these data, nor can such a problem be documented for secondary roads or for city or town streets.

If the speed limit differential between school buses and other traffic was a factor in these accidents, a preponderance of sideswipe and rear-end accidents would be expected. Indeed, Table 3 indicates that these maneuvers were involved in 80 percent of Interstate highway accidents involving school buses but accounted for only 40 percent of the crashes on primary roads and 24 percent of those on secondary roads.

In an attempt to determine whether the July 1988 increase in the speed limit for passenger vehicles had an impact on the incidence of school bus accidents on Interstate highways, an analysis was conducted of all school bus accidents on those highways between September 1985 and May 1989. Most of these accidents, both before and after the change in the maximum speed limit, resulted in no injuries.

If the increased speed limit differential between passenger vehicles and school buses had increased the probability of accidents, the number of accidents would have been expected to increase. This situation appears to be the case; 10 accidents occurred during the three previous school years, compared with 7 during the school year after the change was made. These numbers, however, are small and could reflect random fluctuations rather than a trend. However, if accident probability did increase, perhaps accident characteristics mirror this change. Assuming that the probability did increase, it would be expected that more accidents would occur in which vehicles other than school buses were traveling faster than the buses and either rear-ended or sideswiped them. Table 4 indicates that before the change in the rural Interstate highway speed limit for passenger vehicles, the most common type of collision involving a vehicle and a school bus was a sideswipe or angle-type accident. The next most common type of collision involved the other vehicles rear-ending school buses. These two accident types did not increase after the speed limit differential was increased. In fact, after the speed limit change, school buses more often rear-ended other vehicles rather than the other way around. Hence, although there were relatively few rural Interstate crashes both before and after the speed limit for passenger vehicles was increased on rural Interstate highways, configuration data do not support arguments that the increased speed limit differential resulted in more crashes.

Finally, a major fear about a speed limit differential for school buses is that very large vehicles might strike a bus from behind, causing serious injuries to students. Table 5 indicates that almost all Interstate highway school bus accidents involved cars, and none involved large trucks. In fact, the only large vehicles to strike school buses in the before and after periods were other school buses.

Opinion Surveys

As explained previously, each of the three groups surveyed (pupil transportation administrators, school bus drivers, and police agencies and other special interest groups) was asked several questions in common and then a number of questions specific to the interests and expertise of the individual group. The first common question concerned the ideal maximum speed limit for school buses on Interstate highways.

A majority of each of the three groups surveyed supported a 55-mph speed limit on rural Interstate highways but preferred 45 mph for urban Interstate highways. The majority of school bus drivers believed that their buses could travel safely on rural Interstate highways at 55 mph but that their vehicle could not adequately climb hills at that speed. The drivers' perceptions supported the hypothesis developed from the speed survey that a speed governor tends to limit travel speeds to about 45 mph. Both administrators and bus drivers were opposed to prohibiting school buses from traveling on Interstate highways.

DISCUSSION OF RESULTS

The guiding principle of this study was to identify speed limit policies that would ensure the safe travel of students on school buses. A synthesis of the data suggests clear directions for some policy issues, but directions for other issues remain unclear. Obviously, it is impossible to eliminate the risk of injury

TABLE 3 SCHOOL BUS CRASH CONFIGURATION BY LOCATION

Location	Sideswipe—Same Direction (%)	Rear End—Vehicle Striking Bus (%)	Other Collision (%)
Interstate Highway	4 (40.0)	4 (40.0)	2 (20.0)
Primary Highway	32 (10.7)	87 (29.2)	179 (60.1)
Secondary Road	69 (7.0)	173 (17.5)	749 (75.5)
City/Town Street	129 (12.7)	207 (20.5)	676 (66.8)
School Facility	21 (17.1)	9 (7.3)	93 (75.6)

TABLE 4 INTERSTATE HIGHWAY SCHOOL BUS CRASH CONFIGURATION, SEPTEMBER 1985 TO MAY 1989

Configuration	9/85 - 5/88 Before Change	6/88 - 5/89 After Change
Other Vehicle Rear-Ends Bus	3	1
Bus Rear-Ends Other Vehicle	0	4
Sideswipe or Angle	4	1
Bus Strikes Bus	1	1
Other Collision	2	0

or death in school bus travel; however, compared with other types of travel, school bus travel is extremely safe.

Placing school buses on high-speed highways comes with its own risks. On the one hand, the faster any vehicle travels, the greater the potential risk for injury or death if the vehicle in involved in a crash. A vehicle traveling far slower than the prevailing speed on a highway, however, is at a higher risk for being involved in a crash. Thus, neither slower nor faster is always better. Each has its own advantages and disadvantages.

All of the options concerning school bus speed limits were considered carefully, and several were eliminated. The option of allowing school buses to travel at a maximum of 60 or 65 mph on rural Interstate highways was eliminated. Although 22 states have a maximum speed limit of 65 mph and 1 has a maximum of 60 mph, these limits would be more than the limit for trucks in Virginia. A higher speed limit for school buses than for trucks would not reflect, and would be inconsistent with, the rationale for establishing a truck speed limit of 55 mph. A speed limit of 65 mph would also be inconsistent with Virginia's tradition of establishing a speed limit for school buses lower than that for passenger vehicles.

The option of lowering the current 45-mph maximum speed limit for school buses was also eliminated. Virginia is already among four states with a 45-mph maximum speed limit, and only South Carolina has a lower maximum speed limit. There is no indication from accident and speed data that Virginia's school buses are currently traveling too fast. Thus, lowering the maximum speed limit would likely increase crash risk without reducing potential crash severity.

Option 1: Increase School Bus Speed Limit to 55 mph

One option is to permit school buses to travel 55 mph on rural Interstate highways only or on both urban and rural Interstate highways. An advantage of this option is that the change would result in more uniformity with other states' statutes or regulations. Further, on the basis of previous research, the accident probability should decrease because the differential between passenger vehicle, truck, and school bus speed limits would be reduced. Virginia has traditionally had a speed limit differential between school buses and passenger vehicles, and this change would be in keeping with this tradition. The truck and school bus speed differential would be eliminated, but the passenger vehicle and school bus differential of 10 mph would remain. This alternative was strongly supported for rural Interstate highways by administrators, bus drivers, and police agencies and other special interest groups.

This option also has a number of disadvantages. For example, because school buses would be traveling faster than they currently are, accident severity would likely increase. This increase would probably be most severe in accidents involving school buses and large trucks. Having two tiers of speed limits would still allow a lot of speed variance in the system, thereby theoretically increasing crash probability over that for a uniform speed limit.

One clear-cut problem associated with increasing the speed limit for school buses on Interstate highways involves speed governors. Increasing the speed limit to 55 mph would also necessitate increasing the maximum speed allowed by a governor. Currently a governor is set at 45 mph, which is the maximum speed school buses can legally travel on other systems. If a governor is set at 55 mph, however, it will allow speeds higher than the maximum limits set for school buses on all other roadway systems. Thus, the effectiveness of a governor in controlling speeds on other systems would be reduced, and school bus speeds might increase on other systems, particularly primary highways.

Option 2: Increase School Bus Speed Limit to 50 mph

Another option would be to increase the maximum speed limit for school buses to 50 mph on rural Interstate highways only or on both urban and rural Interstate highways. Increas-

TABLE 5 OTHER VEHICLE TYPES INVOLVED IN INTERSTATE HIGHWAY SCHOOL BUS CRASHES, SEPTEMBER 1985 TO MAY 1989

Vehicle Type	Before Change (%)	After Change (%)
Car	8 (80.0)	5 (71.4)
Pickup Truck	1 (10.0)	1 (14.3)
Large Truck	0	0
School Bus	1 (10.0)	1 (14.3)

ing the limit to 50 mph on Interstate highways is supported by speed theory in that it would reduce the speed limit differential and therefore should reduce accident probability. However, this same theory lends more support to a 55-mph speed limit for school buses.

A disadvantage of raising the speed limit for school buses to 50 mph is that accident severity should increase, but not as much as it would with a 55-mph speed limit. Increasing the school bus speed limit to 50 mph would also reduce the effectiveness of a speed governor on other roadway systems but, again, not as much as with a 55-mph speed limit. In addition, increasing the limit to 50 mph would not eliminate the three-tiered speed limit system for rural Interstate highways.

Option 3: Maintain the School Bus Speed Limit at 45 mph

Another option would be to retain the 45-mph maximum speed limit for school buses on Interstate highways. An advantage of this option is that, according to speed theory, no increase in crash severity would result. Retaining the current speed limit would also maintain the effectiveness of a speed governor on other systems.

A disadvantage of maintaining the current speed limit is that Virginia's policies concerning school bus speed limits are unlike those adopted by most other states. Under this option Virginia would be among those states with the lowest maximum speed limit in the nation. Further, this option would do nothing to mitigate the problem of the 20-mph speed differential, which, according to theory, should result in increased crash probability.

CONCLUSIONS AND RECOMMENDATIONS

Although 45, 50, and 55 mph are all viable options for maximum speed limits for school buses on Virginia's Interstate highways, none of these options can eliminate the risks associated with transporting students. An underlying assumption in research is that, unless there is compelling evidence to indicate that a change is needed, the status quo should be maintained. Although the theoretical accident probability should have increased when the speed limit for most other vehicles was raised on rural Interstate highways, accident data from Virginia's 1 year of experience with the increased differential do not indicate that this increased probability was manifested. Thus, because there were only a few, relatively minor, school bus crashes on the Interstate highways under the current maximum school bus speed limit of 45 mph, and because there was not sufficient evidence to support the hypothesis that school bus accident probability increased when the 65-mph rural Interstate highway speed limit for most vehicles was implemented, the study team concluded that there was not enough compelling evidence to warrant a change in current speed limit policies pertaining to school buses in Virginia. In addition, one positive aspect of retaining the 45-mph Interstate speed limit for school buses involves the effectiveness of a speed governor. Retaining the 45-mph speed limit on Interstate highways precludes raising the limit on a speed governor, thereby preserving the efficacy of the device in contributing to speed limit compliance on other roads.

REFERENCES

1. *Special Report 222: Improving School Bus Safety.* TRB, National Research Council, Washington, D.C., 1989.
2. D. Solomon. *Accidents on Main Rural Highways Related to Speed, Driver and Vehicle.* FHWA, U.S. Department of Transportation, 1968.
3. B. Kelley. *Speed and Car Crashworthiness: Unfixing a Great Gulf.* Insurance Institute for Highway Safety, Washington, D.C., Nov. 1973.
4. P. Johnson, T. Klein, P. Levy, and D. Maxwell. *The Effectiveness of the 55 mph National Maximum Speed Limit.* NHTSA, U.S. Department of Transportation, 1976.
5. R. Michaels and C. Schneider. The Energy Crisis, Characteristics of Traffic Flows and Highway Safety. *Accident Analysis and Prevention*, Vol. 8, 1976.
6. *Speed and Accidents.* Research Triangle Institute, Durham, N.C., June 1976.
7. N. J. Garber and R. Gadirau. *Speed Variance and Its Influence on Accidents.* AAA Foundation for Traffic Safety, Washington, D.C., 1988.
8. E. Hauer. Accidents, Overtaking and Speed Control. *Accident Analysis and Prevention*, Vol. 3, 1971.

Publication of this paper sponsored by Committee on Traffic Law Enforcement.

Impact of 65-mph Speed Limit on Virginia's Rural Interstate Highways Through 1989

JACK D. JERNIGAN AND CHERYL W. LYNN

In April 1987, Congress passed the Surface Transportation and Uniform Relocation Assistance Act, which permitted states to raise their maximum speed limit on rural Interstate highways to 65 mph. Since then, 40 states (including Virginia) have adopted a 65-mph maximum speed limit. Virginia's 65-mph speed limit became effective for passenger cars on July 1, 1988, and for commercial buses on July 1, 1989. The findings presented summarize 18 months of experience with the 65-mph speed limit in Virginia. Fatal crashes and fatalities increased on average more in Virginia than in other states that raised their maximum speed limit to 65 mph. The change in the maximum speed limit may have caused these increases. However, other factors cannot be ruled out. For instance, weather conditions, changes in traffic volume, trip type, or vehicle mix could account for some of the increase. Thus, although increases in speed occurred, the change in the maximum speed limit may not account for the increases in fatal crashes and fatalities.

In 1974, Congress established the 55-mph national maximum speed limit (NMSL) as an energy conservation measure in response to the Arab oil embargo. In addition to reducing the consumption of fuel oil, the 55-mph NMSL helped to reduce drastically the number of deaths on the nation's highways (*1*). Before establishment of the 55-mph NMSL, higher speed limits were common on the nation's Interstate, primary, and unposted secondary highways. For more than a decade the 55-mph NMSL remained, but, as energy constraints lessened and fuel prices decreased in the 1980s, public pressure began to mount for a lessening of federal control. One federal policy, the federal speed compliance-monitoring program, dictated that states maintain a 50 percent minimum level of compliance with the 55-mph speed limit or risk having as much as 10 percent of their federal-aid highway funds impounded. Many states were concerned that their highway funding was in jeopardy because the federal compliance-monitoring program included rural Interstate highways, for which the level of compliance was the lowest.

In response to public pressure for higher speed limits and pressure from the states to avoid the threat of losing highway funds, Congress passed the Surface Transportation and Uniform Relocation Assistance Act (STURAA) in April 1987. The act included a provision to allow states to increase the maximum limit to 65 mph, without penalty, for Interstate highways outside urbanized areas with a population of 50,000 or more.

Virginia Transportation Research Council, Box 3817 University Station, Charlottesville, Va. 22903-0817.

During 1987, 38 states increased the maximum speed limit on at least part of their rural Interstates. In 1988, Georgia became the 39th state to increase the rural Interstate speed limit to 65 mph. On July 1, 1988, the speed limit on most of Virginia's rural Interstates was raised from 55 to 65 mph for passenger vehicles but remained at 55 mph for commercial buses and large trucks. A year later, however, the speed limit for commercial buses was raised to 65 mph, although the truck speed limit remained at 55 mph.

METHODOLOGY

The focus of the following paragraphs is on changes in travel speeds, fatal crashes, and fatalities that occurred on rural Interstates after implementation of the 65-mph speed limit. Data for urban Interstate highways are compared with data for rural Interstates to determine whether similar patterns emerge for these highways, even though the urban Interstates have a 55-mph maximum speed limit.

In Virginia speed data were collected from the permanent sites used in the federal compliance-monitoring program. Although the federal government no longer requires that speeds be monitored on Interstate highways posted at 65 mph, the Virginia Department of Transportation (VDOT) elected to continue to collect data at these stations. Speed data for other states were solicited by contacting the state agency responsible for conducting the federal compliance-monitoring survey and were supplemented by data from published sources (*2–4*). About half of the states with a 65-mph rural Interstate speed limit no longer monitor speeds on these highways; thus, the speed data for other states' rural Interstates are based on a self-selected sample. Most states were able to provide speed data for urban Interstates because the federal compliance-monitoring program is still in effect for highways with a 55-mph speed limit. Generally, speed data for the spring quarter (April to June) for the years between 1986 and 1989 are compared here. However, because of the limited amount of data, a few states were compared across other time periods.

In addition to the speed data collection, daytime radar speed surveys were conducted before and after the increase in Virginia's rural Interstate speed limit. The radar survey allowed the study team to distinguish between the speeds of cars and trucks, which have different speed limits.

Because all 40 states that changed the rural Interstate speed limit to 65 mph did so on a date other than January 1, all states had a year of transition during which the speed limit changed. The study team elected to compare the calendar year immediately before the transition year with the calendar year following the transition year. Thus, with the exception of Georgia, which like Virginia did not raise its speed limit until 1988, the "before" year in the section on fatal crashes and fatalities is 1986 and the "after" year is 1988. In Virginia and Georgia, the before year is 1987 and the after year is 1989. Because conditions other than the maximum speed limit also change from year to year, great caution must be used in comparing data from different years.

Finally, a regression model was calculated, estimating the number of rural Interstate fatalities from annual average speed and vehicle-miles of travel (VMT). The years 1966–1987 were used as the baseline data for the model, and projections were made for 1988 and 1989 on the basis of this model.

SPEEDS AND CRASHES ON INTERSTATE HIGHWAYS

Speeds

Rural Interstates

Actual speeds on Virginia's rural Interstates increased after the implementation of the 65-mph speed limit, but substantially less than the 10-mph increase in the legal limit. In Virginia, as in many other states, speeds on rural Interstates increased between 1986 and 1987 as the passage of the STURAA became inevitable. In the spring of 1986 the average speed traveled on Virginia's rural Interstates was 56.3 mph, and the 85th-percentile speed was 62.0 mph. During the spring of 1987 the average speed traveled on rural Interstates had increased to 59.9 mph and had further increased to 63.5 mph by the spring of 1989. However, as can be seen in Figure 1, the average speed was still lower than 65 mph. The 85th-percentile speed (the speed at or below which 85 percent of vehicles travel) was 65 mph in the spring of 1987 but had increased to 70 mph during the same time period in 1989.

These speed figures may seem lower than the speeds usually experienced while driving on rural Interstates; however, these data represent averages based on 24-hr surveys conducted in varying weather and traffic conditions. At night and during inclement weather, people tend to drive more slowly than on sunny days.

Subjective experience is also not a reliable reflection of actual travel speeds. For instance, some reporters have traveled on rural Interstates (usually on a sunny day when speeds are highest) with the vehicle's cruise control set at 65 mph and have counted the number of vehicles passing their car and the number of vehicles they pass. In this so-called experiment, the only vehicles the reporters would encounter were those traveling substantially faster or slower than they were traveling. Thus, the reporters would never encounter vehicles traveling at a similar speed; that is, they would not catch up with vehicles traveling at similar speeds and vice versa. Hence, surveys such as these tend to detect only abnormal rather than normal travel speeds at a time and under conditions during which speeds are generally at their highest.

Table 1 indicates how average and 85th-percentile speeds on Virginia's rural Interstates compare with the mean of those speeds for other states. Only 10 states do not have a 65-mph speed limit on rural Interstates, but only 7 of those states have more than a handful of rural Interstate miles. The sample of states retaining the 55-mph limit had an average rural Interstate speed of 58.9 mph in 1986 and 61.3 mph in 1989. The 85th-percentile speed in these states increased from 65.9 mph in 1986 to 68.2 mph in 1989. Thus, both average and 85th-percentile speeds were up 2.4 and 2.3 mph, respectively, in states that did not increase the rural Interstate speed limit.

In the sample of states other than Virginia that increased the rural Interstate speed limit, the average speed in 1986 was 60.7 mph, and the 85th-percentile speed was 66.7 mph. By 1989 the speed on rural Interstates posted at 65 mph had increased to 64.4 mph, and the 85th-percentile speed had increased to 70.9 mph. Thus, in states that increased the speed limit, the average rural Interstate speed was up by 3.7 mph, and the 85th-percentile speed was up by 4.2 mph.

In Virginia, the average rural Interstate speed in the spring of 1986 was 56.3 mph, and the 85th-percentile speed was 62.0 mph. These speeds had increased to 63.5 and 70.0 mph, re-

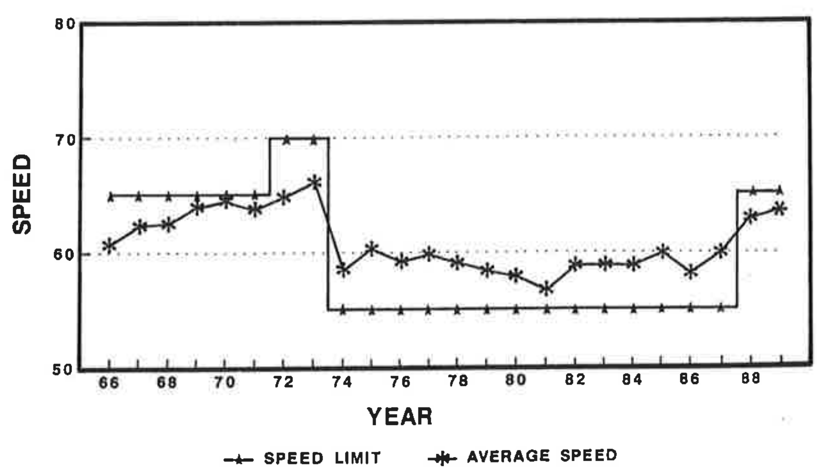


FIGURE 1 Rural Interstate highway average speed versus speed limit in Virginia, 1966–1989.

TABLE 1 AVERAGE AND 85th PERCENTILE RURAL INTERSTATE SPEEDS

	Speed, mph					
	States with 55-mph Rural Interstate Speed Limit (n=6)		States with 65-mph Rural Interstate Speed Limit (n=19)		Virginia	
	Avg.	85th Percentile	Avg.	85th Percentile	Avg.	85th Percentile
1986	58.9	65.9	60.7	66.7	56.3	62.0
1987	59.7	66.3	61.7	68.7	59.9	65.0
1988	60.6	67.8	63.1	69.2	60.1	66.0
1989	61.3	68.2	64.4	70.9	63.5	70.0
Change 86–89	+2.4	+2.3	+3.7	+4.2	+7.2	+8.0

NOTE: States not collecting or submitting data are excluded from this table.

spectively, by 1989—an increase of 7.2 mph in the average speed and 8.0 mph in the 85th-percentile speed. However, even though Virginia's increase in speeds was greater than that of other states, the actual average and 85th-percentile speeds on Virginia's rural Interstates remained lower than those for other states with a 65-mph rural Interstate speed limit. Further, of the 19 other states in the sample with a 65-mph rural Interstate speed limit, only 4 reported lower average speeds than Virginia and only 6 reported lower 85th-percentile speeds.

On Virginia's rural Interstates average speeds were only 3.6 mph higher and 85th-percentile speeds 5.0 mph higher in 1989 than in 1987, the year Congress passed the STURAA. Virginia's rural Interstate speeds increased between the spring of 1986 and 1987 by 3.6 mph for the average speed and 3.0 mph for the 85th-percentile speed. Therefore, even without action by the general assembly, travel speed had increased on Virginia's rural Interstates.

Urban Interstates

Like rural Interstate speeds, urban Interstate speeds have increased, but not by as much. Table 2 shows that, in the sample of states in which both urban and rural Interstate speed limits are 55 mph, the average urban Interstate speed was 59.1 mph in 1986 and 60.4 mph in 1989. The 85th-percentile speed in these states increased from 65.6 mph in 1986 to 67.8 mph in 1989. Thus, between 1986 and 1989 in states that did not change the speed limit on rural Interstates, the average urban Interstate speed increased 1.3 mph, and the 85th-percentile speed increased 2.2 mph.

In states that increased the rural Interstate speed limit, the average urban Interstate speed was 58.0 mph in 1986 and 59.4 in 1989, and the 85th-percentile speed increased from 64.2 mph in 1986 to 66.1 mph in 1989. Hence, the average urban Interstate speed increased 1.4 mph and the 85th-percentile speed 1.9 mph. Even though the increase in the urban Interstate speed was similar in states that increased the rural Interstate speed limit and in those that did not, the average and 85th-percentile urban Interstate speeds were lower in the states that had a 65-mph rural Interstate speed limit.

In Virginia the increases in average and 85th-percentile speeds on urban Interstates were greater than those in other states. The average speed went from 53.5 to 58.7 mph, and the 85th-percentile speed increased from 61.0 to 68.0 mph. Hence, in Virginia the urban Interstate average speed increased 5.2 mph and the 85th-percentile speed 7.0 mph. However, the average speed in Virginia was lower than that for other states, although the 85th-percentile speed was higher.

The distribution of speeds also changed. Speed variance is a measure of the distribution of speeds on a section of highway; low speed variance indicates a more uniform travel speed than higher speed variance. Speed theory suggests that, as speed variance increases, the number of interactions between vehicles increases, thereby increasing the overall accident potential. Thus, with the introduction of a differential speed system (in which the maximum speed limit for trucks remained at 55 mph but the limit for passenger cars was increased), it was anticipated that speed variance would also increase. Speed variance, in fact, did increase in Virginia (see Table 3).

At survey sites on the rural Interstates, speed variance increased an average of 36.4 percent after the speed limit was raised. This result is the opposite of what occurred when

TABLE 2 AVERAGE AND 85TH PERCENTILE URBAN INTERSTATE SPEEDS

	Speed, mph					
	States with 55-mph Rural Interstate Speed Limit (n=7)		States with 65-mph Rural Interstate Speed Limit (n=27)		Virginia	
	Avg.	85th Percentile	Avg.	85th Percentile	Avg.	85th Percentile
1986	59.1	65.6	58.0	64.2	53.5	61.0
1987	58.5	65.7	58.1	65.1	53.7	63.0
1988	59.0	66.5	58.7	65.0	59.5	66.0
1989	60.4	67.8	59.4	66.1	58.7	68.0
Change 86–89	+1.3	+2.2	+1.4	+1.9	+5.2	+7.0

NOTE: States not collecting or submitting data are excluded from this table.

TABLE 3 RADAR SURVEY OF SPEEDS ON VIRGINIA'S INTERSTATE HIGHWAYS, SPRING 1988 VERSUS FALL 1989

Site Location		Before (mph) Mean	Before (mph) Variance	After (mph) Mean	After (mph) Variance	Percent Change Mean	Percent Change Variance
Urban							
I-64	Cars	60.51	12.67	60.79	17.41	0.46	37.37
	Trucks	57.95	11.90	58.44	13.20	0.85	10.90
	Total	59.59	13.99	60.17	17.97	0.97	28.47
I-64	Cars	60.70	14.82	64.15	19.35	5.68	30.54
	Trucks	58.18	24.21	61.93	11.45	6.45	-52.70
	Total	59.96	18.66	63.52	17.99	5.94	-3.60
I-95	Cars	62.18	19.89	64.05	29.65	3.01	49.06
	Trucks	58.68	13.91	61.12	16.95	4.16	21.83
	Total	60.63	20.25	63.23	27.49	4.29	35.75
I-95	Cars	59.22	15.05	64.10	24.53	8.24	62.94
	Trucks	57.74	12.25	59.48	22.42	3.01	83.02
	Total	58.70	14.52	62.40	28.54	6.30	96.61
Rural							
I-64	Cars	62.54	19.89	68.23	21.21	9.10	6.63
	Trucks	59.26	15.37	61.91	12.42	4.47	-19.17
	Total	61.73	20.70	66.10	27.17	7.08	31.24
I-85	Cars	61.37	17.31	68.51	16.82	11.63	-2.81
	Trucks	57.03	14.90	62.60	20.35	9.77	36.58
	Total	60.30	19.98	66.88	25.60	10.91	28.12
I-64	Cars	62.54	19.89	67.79	16.37	8.39	-17.70
	Trucks	59.26	15.37	61.89	11.43	4.44	-25.62
	Total	61.73	20.70	65.79	22.44	6.58	8.39
I-95	Cars	61.76	19.62	68.67	14.81	11.19	-24.53
	Trucks	59.55	11.76	68.87	15.06	15.65	28.01
	Total	61.24	18.66	65.61	29.40	7.14	57.54
I-95	Cars	62.00	19.71	68.19	18.17	9.98	-7.83
	Trucks	59.69	11.22	61.63	17.54	3.25	56.29
	Total	61.13	17.64	65.90	27.66	7.80	56.80

Interstate speed limits were lowered to 55 mph in the early 1970s. At that time, as mean Interstate speeds decreased, so did speed variance. At urban Interstate survey sites, speed variance also increased by 39.3 percent. Thus, speed variance increased both on rural and on urban Interstates in Virginia after the implementation of the 65-mph speed limit on rural Interstates.

Crashes

Fatal Crashes and Fatalities on Interstate Highways

Generally, fatal crashes and fatalities both increased nationwide after implementation of the 65-mph speed limit. However, these increases may result, in part, from other factors or normal fluctuations in the data.

The data in Table 4 indicate that fatal crashes on rural Interstates in states that did not increase the speed limit increased 17.3 percent and that fatalities increased 11.1 percent in the calendar year before and after the change in the speed limit. In states that increased the speed limit, fatal crashes on rural Interstates increased 32.2 percent and fatalities increased 34.7 percent in those years. Thus, fatalities and fatal crashes on rural Interstates increased more in states that raised their speed limit than in those that did not.

In Virginia fatal crashes increased from 40 in 1987 to 59 in 1989, and fatalities increased from 44 to 63. Thus, Virginia had a 47.5 percent increase in fatal crashes and a 43.2 percent increase in fatalities.

Table 5 indicates that, on urban Interstates in states not increasing the maximum speed limit to 65 mph, fatal crashes and fatalities rose by 25.9 and 27.4 percent, respectively, between 1986 and 1988. Increases in urban Interstate crashes in states raising the rural Interstate speed limit were much lower, with fatal crashes increasing only 4.9 percent and fatalities only 4.1 percent. In Virginia between 1987 and 1989 (1 year before and 1 year after the speed limit increase), fatalities increased 4.9 percent and fatal crashes increased 4.3 percent. Like other states that increased the rural Interstate speed limit, Virginia's percentage increases in fatal crashes and fatalities on urban Interstates were less than those in states that retained the 55-mph speed limit for rural Interstates.

Characteristics of Fatal Crashes on Rural Interstates

Tables 6 and 7 present the breakdown of fatal crashes and fatalities in Virginia by month for 1985 through 1989. Com-

TABLE 4 FATAL CRASHES AND FATALITIES ON RURAL INTERSTATES

	55-mph States (86-88) (n = 8)	65-mph States (86-88) (n = 38)	Virginia (87-89)
Changes in fatal crashes	+17.3%	+32.2%	+47.5%
Changes in fatalities	+11.1%	+34.7%	+43.2%

States not collecting or submitting data are excluded from these tables.

TABLE 5 FATAL CRASHES AND FATALITIES ON URBAN INTERSTATES

	55-mph States (86-88) (n = 9)	65-mph States (86-88) (n = 37)	Virginia (87-90)
Changes in fatal crashes	+25.9%	+4.9%	+4.9%
Changes in fatalities	+27.4%	+4.1%	+4.3%

States not collecting or submitting data are excluded from these tables.

TABLE 6 FATAL CRASHES ON VIRGINIA'S RURAL INTERSTATES BY MONTH, 1985–1989

	1985	1986	1987	1988	1989	85-87 Avg.	89 Diff.
Jan.	2	3	3	3	3	2.7	+0.3
Feb.	1	1	4	4	4	2.0	+2.0
Mar.	3	5	3	3	2	3.7	-1.7
Apr.	4	1	1	5	2	2.0	0.0
May	2	3	4	5	9	3.0	+6.0
June	5	3	2	6	6	3.3	+2.7
July	6	2	4	5	6	4.0	+2.0
Aug.	7	6	5	6	5	6.0	-1.0
Sept.	5	6	5	5	8	5.3	+2.7
Oct.	6	7	2	12	9	5.0	+4.0
Nov.	5	1	7	6	4	4.3	-0.3
Dec.	4	2	0	5	2	2.0	0.0
TOTALS	50	40	40	65	59	43.3	+15.7

TABLE 7 FATALITIES ON VIRGINIA'S RURAL INTERSTATES BY MONTH, 1985–1989

	1985	1986	1987	1988	1989	85-87 Avg.	89 Diff.
Jan.	2	4	3	3	3	3.0	0.0
Feb.	1	1	5	5	4	2.3	+1.7
Mar.	4	5	3	3	2	4.0	-2.0
Apr.	5	1	1	6	2	2.3	+0.3
May	4	3	4	5	10	3.7	+6.3
June	6	3	3	6	7	4.0	+3.0
July	6	3	6	7	6	5.0	+1.0
Aug.	8	7	5	8	5	6.7	-1.7
Sept.	5	7	5	7	8	5.7	+2.3
Oct.	8	7	2	13	11	5.7	+5.3
Nov.	6	1	7	8	4	4.7	-0.7
Dec.	4	2	0	7	2	2.0	0.0
TOTALS	59	44	44	78	63	49.0	+14.0

pared with the average in the 3 years before implementation of the higher speed limit, a substantial portion of the 1989 increase in fatal crashes and fatalities occurred during May and October, and there were several months in which there was either no change or a reduction. Hence, increases in fatal crashes and fatalities were not evenly distributed across the year.

Various characteristics of fatal crashes occurring on rural Interstates are presented in Table 8. The percentage of all crashes that involved speeding as a contributing factor remained relatively constant between 1987 and 1989 compared with other characteristics. (However, the speeds defined as excessive changed when the speed limit was increased from 55 to 65 mph.) The categories expected to increase because of the differential speed limit for cars and trucks (rear end, sideswipe, and truck involved) all declined over this time period, as did the percentage of crashes involving pedestrians and alcohol. The only categories to increase in representation between 1987 and 1989 were accidents involving running off the road and driving the wrong way.

Table 9 presents the distribution of fatal crashes on Virginia's rural Interstates by route. These data show that, in each year, the majority of all rural Interstate fatal crashes occurred on I-81 and I-95, which is expected considering the length of these roads and the volume of traffic on them. Fatal crashes on I-81 and I-95 each increased by three to four crashes in 1989 compared with the 3-year average from 1985 to 1987.

Between 1987 and 1989, fatal crashes increased disproportionately on I-77 and I-85 compared with other Interstate routes; however, such a disproportionate increase was not apparent in 1988. On I-77, the 3-year average was 2.0 fatal crashes per year, but in 1989 there were 5 fatal crashes on this route. Likewise, on I-85, where the 3-year average was 1.7 fatal crashes per year, there were 7 fatal crashes in 1989.

Generally, the data in Table 9 indicate that the increase in fatal crashes was not route-specific; that is, part of the overall increase in fatal crashes was distributed across most of the routes. Although there are insufficient data to determine whether the disproportionate increases in fatal crashes on I-77 and I-85 represent more than normal fluctuations of the data, the study team will continue to monitor these routes closely.

Relationship Between Average Speed and Fatalities

In order to determine how average speed is related to the number of fatalities on Virginia's rural Interstates, a multiple regression analysis was conducted using data from 1966 to 1987. Table 10 indicates that, after controlling for VMT, there

TABLE 8 CONFIGURATION OF FATAL CRASHES ON RURAL INTERSTATES

	1987 (% of all crashes)	1988 (% of all crashes)	1989 (% of all crashes)
Run off road	30 (75.0)	39 (60.0)	50 (84.7)
Rear end	10 (25.0)	12 (18.5)	8 (13.6)
Sideswipe (same direction)	5 (12.5)	8 (12.3)	6 (10.2)
Pedestrian	6 (15.0)	7 (10.8)	5 (8.5)
Wrong way	1 (2.5)	8 (12.3)	3 (5.1)
Truck involved	27 (67.5)	20 (30.8)	10 (16.9)
Alcohol	8 (20.0)	14 (21.6)	9 (15.3)
Speeding	8 (20.0)	16 (24.6)	13 (22.0)
TOTAL CRASHES	40	65	59

Note: Individual categories add up to more than the total because more than one factor could be involved in one crash (e.g., an alcohol-related run-off-road crash).

TABLE 9 FATAL CRASHES ON VIRGINIA'S RURAL INTERSTATES BY ROUTE

Route	1985	1986	1987	1988	1989	85-87 Avg.	89 Change
64	4	8	7	9	7	6.3	+0.7
66	2	3	3	6	2	2.7	-0.7
77	3	1	2	2	5	2.0	+3.0
81	17	14	16	26	19	15.7	+3.3
85	1	3	1	3	7	1.7	+5.3
95	21	11	11	19	18	14.3	+3.7
295	2	0	0	0	1	0.7	+0.3
TOTAL	50	40	40	65	59	43.3	+15.7

TABLE 10 REGRESSION ANALYSIS OF VMT AND AVERAGE SPEED ON TRAFFIC FATALITIES

Variable	b	Standard Error	Sig.
Average speed	3.9107	.980	0.001
VMT (millions)	-0.0069	0.0028	0.023
(Constant)	-143.5876	66.5638	0.044

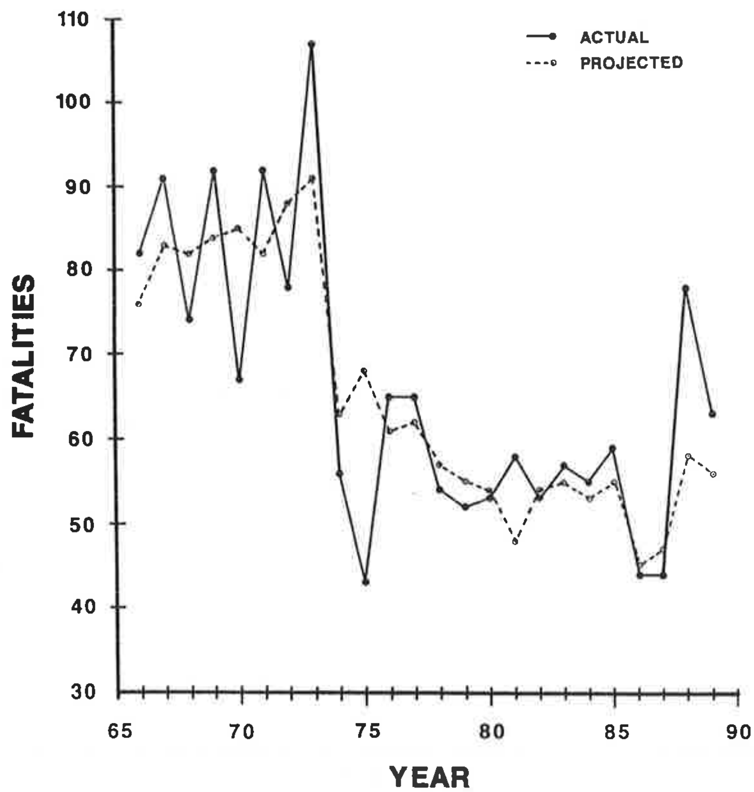


FIGURE 2 Actual and projected fatalities on Virginia's rural Interstate highways, 1966–1989.

was a significant positive relationship between the annual average speed on Virginia's rural Interstates and fatalities on those highways in a given year. Thus, as average speed increase, fatalities increase.

There is a significant negative correlation between VMT and fatalities. As traffic has increased, fatalities have decreased—the opposite of what would normally be expected. A possible explanation for this relationship is that VMT has increased relatively steadily over the years and is therefore closely associated with annual improvements in highway and vehicle safety. At least since the late 1960s, highways and vehicles have been designed to reduce fatalities in crashes. Hence, increases in fatalities that might be expected with increased VMT may be offset by the effects of annual improvements in highway and vehicle safety. (Figure 2 shows the actual number of annual fatalities on rural Interstates compared with the projected numbers of fatalities using only average speed and VMT as predictor variables.)

The regression model estimates that a 1-mph increase in the annual average speed on Virginia's rural Interstates is associated with an increase of approximately 4 deaths on those highways, all other factors being equal. On the basis of the 95 percent confidence range, an increase of 1 mph in the average speed would be associated with between 2 and 6 deaths on Virginia's rural Interstates. Thus, if all other factors were held constant, this model would predict that the 3.6-mph increase in the average speed on Virginia's rural Interstates from 1987 to 1989 would have been associated with an increase of between 7 and 22 rural Interstate fatalities. In fact, there was an increase of 19 fatalities. However, not only did the average speed change, but between 1987 and 1989 VMT also increased. The regression model estimates that an increase of 1 billion VMT on rural Interstate highways is associated with a decrease or approximately 7 deaths on rural Interstates annually. The 95 percent confidence range estimates that an increase of 1 billion VMT has been associated with between 4 and 10 fewer deaths on Virginia's rural Interstates. Thus, given the increase in average travel speeds and VMT, the model predicted that there would be 56 fatalities in 1989—7 fewer than the actual total of 63.

DISCUSSION OF RESULTS

A comprehensive examination of the data indicates that significant increases in the speed of traffic and in the number of fatal crashes and fatalities occurred on Virginia's Interstates in 1989 compared with time periods when the maximum speed limit was 55 mph. It is premature to suggest that the change in the speed limit caused these changes, because adequate time has not passed and sufficient data have not been collected to rule out the influence of other factors.

During the early to mid-1980s Virginia was among the states with the lowest average and 85th-percentile travel speeds on rural Interstates. In 1989 the average and 85th-percentile speeds on Virginia's rural Interstates had become more comparable to those of other states with a 65-mph speed limit than they were when all states were subject to the 55-mph NMSL. In the spring of 1986, a time before a change in national policy had become likely, the average speed on Virginia's rural Interstates was 56.3 mph, and the 85th-percentile speed was 62.0 mph. By the spring of 1989, almost a year after Virginia's rural Interstate speed limit was increased to 65 mph, the average speed had increased to 63.5 mph and the 85th-percentile speed to 70 mph. However, half of the increase in the average speed and 3 mph of the increase in the 85th-percentile speed had already occurred by the spring of 1987—before the general assembly had even considered raising the speed limit on rural Interstates (but after Congress had cleared the path for states to do so).

A regression model of annual average speed and VMT on fatalities indicates that there is a significant positive relationship between average speed and the annual number of fatalities. Thus, it was anticipated that an increase in speeds, which would likely follow an increase in the speed limit, would be associated with an increase in fatalities on Virginia's rural Interstates. In fact, fatal crashes and fatalities did increase in 1988 and 1989 compared with 1986 and 1987. However, 1986 and 1987 were years during which fatalities on Virginia's rural Interstates were at their lowest levels since 1975. Thus, the increase in fatal crashes and fatalities may appear larger when compared with recent years as opposed to comparisons with historical trends.

Speed theory suggests that the speed limit differential for trucks would have increased speed variance on rural Interstates, thereby increasing collisions between passenger vehicles and trucks. A daytime radar survey of passenger vehicle and truck speeds indicated that speed variance increased in 1989 compared with the spring of 1988, when all vehicles had a speed limit of 55 mph. However, accident data indicate that fatal accidents involving trucks did not increase after the speed limit for passenger vehicles was raised to 65 mph.

In conclusion, it appears that there was an increase in average and 85th-percentile speeds, fatal crashes, and fatalities on Virginia's rural Interstates after implementation of the 65-mph speed limit. These increases occurred in most states that raised the rural Interstate speed limit to 65 mph as well as in those that retained a 55-mph speed limit. However, Virginia's increases in fatal crashes and fatalities exceeded those of other states and speeds increased more in Virginia than in other states, thereby closing the gap between average speeds in Virginia and those in other states.

Fatal crashes involving trucks did not increase in Virginia between 1987 and 1989. Also, in Virginia, as in the other states that increased the speed limit on rural Interstates, there was no dramatic increase in fatalities on urban Interstates.

As time passes, the impact of the 65-mph speed limit on Virginia's rural Interstates, exclusive of other factors, will become more apparent. The study team will continue to monitor changes that may be related to the 65-mph speed limit.

REFERENCES

1. *Special Report 204: 55: A Decade of Experience.* TRB, National Research Council, Washington, D.C., 1984.
2. *Highway Safety Performance: 1986: Fatal and Injury Accident Rates on Public Roads in the United States.* Report FHWA-SA-88-007. FHWA, U.S. Department of Transportation, 1988.
3. *Highway Safety Performance: 1987: Fatal and Injury Accident Rates on Public Roads in the United States.* Report FHWA-SA-89-003. FHWA, U.S. Department of Transportation, 1988.
4. S. A. Partyka. *Report to Congress on the Effects of the 65 mph Speed Limit Through 1988.* NHTSA, U.S. Department of Transportation, 1989.

Publication of this paper sponsored by Committee on Traffic Law Enforcement.

Abridgment

Safety Impact of the 65-mph Speed Limit: A Time Series Analysis

RONALD C. PFEFER, WILLIAM W. STENZEL, AND BYOUNG DOO LEE

An examination of the impact of the increase in the speed limit in 1987 to 65 mph on rural Interstate highways in Illinois was conducted. Autoregressive integrated moving average (ARIMA) time series intervention analyses using monthly speed data collected by the Illinois Department of Transportation for 15 rural Interstate highway segments were used to examine changes in (a) 85th-percentile speeds and speed variances for cars and large trucks (i.e., vehicles greater than 24 ft in length) and (b) car-truck speed differentials. ARIMA analyses were also used to examine the changes in accident frequency, accident rates, and proportion of car-truck accidents to all accidents. Consideration was given both to all accidents and to fatal and injury accidents only. Using May 1987 as the point of intervention (i.e., the increase in the speed limit), the ARIMA analyses are based on time series consisting of 52 preintervention months and 15 post-intervention months. Results were obtained for four individual highway segments and for all 15 segments collectively. An increase of 4.0 mph in the 85th-percentile speed for cars was detected at the 95 percent confidence level. No change was found in the 85th-percentile speed for trucks or in the speed variance for either cars and trucks. An increase of 2.8 mph in the car-truck speed differential was found at the 80 percent confidence level. Although the frequency of all accidents increased by 14 percent following the speed limit change, no statistically significant change in the frequency of fatal and injury accidents only was found. No change was detected in the rate for all accidents, but an increase of 18.5 percent was detected in the rate for fatal and injury accidents only. Significant reductions were found for both the rate of car-truck fatal and injury accidents and the proportion of car-truck fatal and injury accidents to all fatal and injury accidents.

On April 2, 1987, 13 years after passage of the National Maximum Speed Limit (NMSL), the law was changed to allow speeds up to 65 mph on rural Interstate facilities. The original objective of NMSL was to conserve fuel in response to the Arab oil embargo. With the establishment of the NMSL, the national annual highway fatality frequency fell by more than 9,100 persons. Many gave NMSL credit for this reduction, and safety emerged as the primary argument for its continuation once the oil embargo was no longer in effect.

The enactment of the NMSL, its implementation and monitoring, and its recent modification have occurred within an atmosphere of continuing controversy. Opponents of the law see the action as arbitrary and an imposition on a free society. Economic inefficiency and lost personal time as a result of slower travel are also cited as drawbacks to the law. The safety benefits are discounted as having been brought about by other factors, such as a decline in travel and improvements to vehicles and the roadway system.

However, the preponderance of opinion among policy makers is that the NMSL does save a significant number of lives each year. It is estimated that between 2,000 and 4,000 lives were saved in 1983 (*1*). The same source also estimated that between 2,500 and 4,500 fewer severe injuries occurred in 1983 as a result of the law. Another study (*2*) examined the potential safety impact of the 65-mph speed limit on rural Interstates using two methods for estimating the annual number of fatalities. One method is promoted by the National Safety Council (NSC), whereas the other was developed by the TRB Committee for the Study of Benefits and Costs of the 55-mph NMSL. Both methods predicted an increase in highway fatalities: 200 to 700 deaths per year using the NSC method, and 300 to 450 deaths per year using the TRB method.

The institution of the 65-mph speed limit on the rural Interstate system occurred as support for NMSL was eroding, as evidenced by a trend toward higher speeds in several states. In Illinois, for example, the 85th-percentile speed on freeways in 1987, just before the law change, was 65 mph [Illinois Department of Transportation (IDOT), unpublished]. Although the recommendation of the TRB committee was for general retention of the NMSL, the use of a higher speed limit on rural Interstates was left as an open issue.

Shortly after passage of the revised law, Congress directed IDOT to initiate a series of studies to investigate the impact of higher speed limits on rural Interstate highways. The activities and findings of a study (*3*) conducted to examine the impact of the 65-mph speed limit on rural Interstates in Illinois are described.

METHODOLOGY

Research Questions

The following research questions were considered:

1. Was there a significant change in speed?
2. Was there a concurrent change in crash experience?
3. Was there a relationship between the change in speed and the change in crash experience?
4. Were observed changes in speeds and accidents different for trucks and other vehicles?
5. Was there an increase in speed differentials between trucks and other vehicles?
6. Was there a concurrent change in accidents involving passenger cars and trucks?

R. C. Pfefer and W. W. Stenzel, Research and Development, Traffic Institute, Northwestern University, 405 Church St., Evanston, Ill. 60204. B. D. Lee, Department of Political Science, Northwestern University, 401 Scott Hall, Evanston, Ill. 60201. Current affiliation: Ma-po Gu Kong-Buk Bong 105-86, Seoul, Korea.

7. Was there a relationship between the change in speed differential and the change in car-truck crash experience?

General Approach

The change to 65 mph in Illinois occurred in May 1987 on approximately 1,200 mi of rural Interstate highway. The speed limit for heavy trucks remained at 55 mph on all Interstate facilities. Traffic volume, speed, and crash data from rural interstate highways were provided by IDOT. The data were analyzed using site-specific and aggregate measures of speed and crash experience. Speed data were provided by IDOT, generated from 15 speed-monitoring sites on rural Interstate highways. Use of state reporting systems for speed and accidents allowed an examination of the impact of speed changes on accident severity. More than 90 percent of the accident reports on Interstate highways are submitted by the Illinois State Police, and the overall quality of the reports is considered to be excellent and consistent.

Analytical Method

Changes in speed and accident experience were examined using time series intervention impact analysis on the basis of autoregressive integrated moving average (ARIMA) models (*4,5*). ARIMA modeling enables the analyst to mathematically represent complex time series patterns; intervention impact analysis is a specific application of ARIMA modeling that enables testing for the impact of an intervention after statistically controlling for historic patterns and other known interventions.

The time series quasi-experiment was first proposed by Campbell (*6*) as a way to examine the impact of interventions on social processes. Two requirements for the use of time series quasi-experiments are as follows:

1. The social process under investigation must be operationalized as a time series; and
2. There must be a discrete intervention that divides the time series into two distinct segments, one consisting of all preintervention observations and another consisting of all postintervention observations.

The proposed study of the Illinois speed limit change satisfies both requirements. The social process under investigation (i.e., driving behavior) was operationalized with monthly speed estimates and monthly crash data, and the intervention event (i.e., the speed limit change) was a discrete event that divided each time series into two parts.

Analysis of a time series quasi-experiment focuses on two issues:

1. Did the intervention have an impact on the time series?
2. If an impact occurred, what was the magnitude and form of the impact?

Both issues are examined statistically by comparing the pre- and postintervention segments of the time series.

Data Collection

Speed Data

The IDOT Traffic Operations Division collects speed data at 70 automated data collection sites throughout the state. Of these, 15 are on rural Interstate roads. The speed data are acquired from induction loop detectors buried in the pavement.

The basic data are received as frequency counts within twelve 5-mph speed categories and for two vehicle length classifications. The classification by vehicle length is possible because a pair of induction loops is placed at each site. Vehicles less than or equal to 24 ft in length are tabulated separately from those with lengths greater than 24 ft. This procedure provides an approximate separation of semitrailer trucks from other vehicles. The length defining the two categories was changed from 21 ft at the beginning of 1983. It was assumed that the speeds of vehicles longer than 24 ft were representative of semitrailer truck speeds and that those for shorter vehicles were representative of passenger vehicles (cars).

Speed-monitoring station data are stored on tape in 80-column records for each vehicle type for each hour of each day. A typical tabulation of the data is presented in Table 1. The IDOT Traffic Operations Division provided a monthly aggregation of speed frequencies by speed bin for the period January 1983–July 1988.

A summary of the data collection activities for each candidate site is presented in Table 2. This table includes information about the percent of days and number of months by year for which speed data were recorded. The summary tabulations provided by IDOT for the 15 rural Interstate monitoring stations indicated that there were some months at every site for which no speed data were recorded because of mechanical failure of the speed-monitoring devices.

The raw speed data were used to compute two aggregate measures: (a) 85th-percentile speed to represent the higher end of the speed distribution, and (b) speed variance as a surrogate measure of speed differentials.

Direct measurement of speed differentials between cars and trucks was not possible due to the lack of data for individual pairs of vehicles in a given lane. The mean speed differential of cars and trucks at each speed-monitoring site was selected as a measure of overall speed disparity. It was recognized that this measure is an average and includes vehicles in adjacent lanes as well as those in a vehicle-following situation. The mean speed differential for cars and trucks in a given traffic stream can be determined from the key statistics of their individual speed distributions (*3*).

The calculation of the speed statistics used standard formulas for classified data, with the middle value of each class interval representing the entire class. The exceptions to this rule were the classes at either end of the range. For the lowest interval (i.e., 1 to 30 mph), a value of 28 mph was used. For the highest interval (i.e., 81 mph and higher), a value of 83 mph was used.

TABLE 1 SPEED-MONITORING DATA BY HOUR FOR 1 DAY AT ONE LOCATION, CARS ONLY

Telac Data for Thursday - 12 November 1987

Station Number 1110 — Telac Number 1
Station Type 2 — Polling Group 1
Station Location INT90 MP 82.8: 0.5 MILE SOUTH OF NAGLE A

Eastbound Combined Speeds for Length 1 thru 24 feet for Lane 1 thru 3

Speed (Mph) (LOWER UNIT FOR RANGE)	1	31	36	41	46	51	56	61	66	71	76	81	Total
Time													
01:00	0	0	0	4	33	150	389	307	253	136	42	18	1332
02:00	0	0	0	0	0	8	19	14	10	3	2	2	58
03:00	0	0	0	2	16	68	154	110	61	37	11	3	462
04:00	0	0	0	3	12	82	113	81	52	20	11	5	379
05:00	0	0	1	7	27	95	177	149	81	45	13	3	598
06:00	0	0	0	12	48	207	563	434	448	284	75	21	2092
07:00	623	794	1068	704	417	540	502	347	236	78	10	2	5321
08:00	3519	498	98	20	1	1	0	0	0	0	0	0	4137
09:00	1951	577	752	537	187	36	7	0	0	0	0	2	4049
10:00	586	241	391	422	456	526	676	518	426	190	40	9	4481
11:00	0	0	0	9	55	376	859	854	923	594	124	29	3823
12:00	0	0	1	20	69	280	680	787	851	571	181	38	3478
13:00	0	0	0	3	42	255	576	777	773	621	187	42	3276
14:00	0	0	0	8	72	338	653	852	910	568	178	30	3609
15:00	0	0	1	9	101	361	805	930	1026	555	155	25	3968
16:00	374	88	103	165	348	684	994	891	806	244	31	9	4737
17:00	563	198	158	249	528	854	1037	711	339	60	9	2	4708
18:00	0	0	8	87	570	1134	1435	948	357	55	10	3	4607
19:00	0	0	0	21	302	923	1303	944	655	160	17	4	4329
20:00	0	0	0	1	83	460	909	850	737	322	57	9	3428
21:00	0	0	0	9	44	278	660	658	590	318	82	27	2666
22:00	0	0	0	13	87	277	678	605	610	355	83	24	2732
23:00	0	0	1	1	78	315	699	666	582	281	81	14	2718
24:00	0	1	0	7	54	208	499	475	445	243	62	21	2015
24 Hour Totals	7616	2397	2582	2313	3630	8456	14387	12908	11171	5740	1461	342	73003

Total Vehicles	73003	Total Vehicles over 55 Mph	46009	Percentage of Vehicles over 55 Mph	63.0
Average Speed	54.7	Total Vehicles over 60 Mph	31622	Percentage of Vehicles over 60 Mph	43.3
Median Speed	54.3	Total Vehicles over 65 Mph	18714	Percentage of Vehicles over 65 Mph	25.6
85th %tile Speed	64.5	Total Vehicles over 70 Mph	7543	Percentage of Vehicles over 70 Mph	10.3

The monthly values of the 85th-percentile speeds and speed variances for passenger cars and trucks are shown in Figures 1–4. Monthly estimates of the mean speed differential between passenger cars and trucks are shown in Figure 5.

Accident Data

Accident data were provided by the IDOT Traffic Safety Division. Monthly summaries were provided for each site for January 1983–July 1988. The highway segment lengths associated with each monitoring station were selected on the basis of inspection of maps, personal knowledge of the areas, advice from IDOT representatives, physical data provided by the Office of Planning and Programming, and, in some cases, site visits.

Crash data included frequencies of fatal accidents, personal-injury accidents, and property-damage-only accidents. These data were tabulated separately for accidents involving semitrailer trucks and those not involving semitrailer trucks, as well as accidents involving passenger cars and semitrailer trucks moving in the same direction. The separation of accidents related to semitrailer trucks relates as closely as possible to the vehicle dichotomy resulting from the speed measurement data.

Monthly accident rates were calculated using estimates of vehicle-miles of travel (VMT) provided by IDOT for all roadway sections with the higher speed limit.

Biannual average daily traffic (ADT) estimates for car traffic volume for each segment were aggregated and adjusted using seasonality factors provided by IDOT to yield traffic volumes expressed in millions of VMT for each month for the period January 1983–July 1988.

Monthly VMT values for nontrucks were derived by subtracting the estimated daily VMT values for trucks from the daily VMT values for all vehicles and multiplying the difference by a seasonal factor and the number of days in the month. Monthly VMT values for trucks were obtained by multiplying the estimated daily VMT by the number of days in the month.

Two monthly frequencies of accidents, supplied by IDOT, were used to calculate accident rates:

1. All accidents, and
2. Only accidents involving a car and a truck moving in the same direction.

These frequencies were divided by the monthly VMT figures to obtain time series for accident rates on 65-mph rural Interstate segments.

TABLE 2 SUMMARY OF DATA COLLECTION ACTIVITIES FOR 15 RURAL INTERSTATE SPEED-MONITORING SITES

STATION NUMBER	LOCATION INTERSTATE NUMBER	MILE POST	NO. MILES	DIRECTION FROM	CROSSING FACILITY
181	55	240	0.5	N	US 6
303	80	85.1	4.3	W	IL 29
305	55	231.5	0.5	N	BRACEVILLE RD
377	57	273.1	0.9	S	BUCKLEY OVERHD
479	74	75.8	1.3	E	BRIMFIELD OVERHD
576	74	178	0.9	W	I-57
577	72	66.4	0.8	E	WHITEHEALTH RD.
676	55	78.1	1.8	S	DIVERNON OVERHD
677	72	15.2	4	W	ILLIOPOLIS OVERHD
703	70	89.8	2	W	I-57
707	57	168	4.9	N	I-70
775	64	68.5	0.7	W	WOODLAWN RD.
802	55	23.1	0.7	N	IL 43
975	57	29.3	0.3	S	US 45
976	24	36.3	0.7	N	US45

PERCENT REPORTING DAYS AND NUMBER OF MONTHS FROM SPEED MONITORING STATION

STATION NUMBER	1983 % DAYS	1983 NO. MONTHS	1984 % DAYS	1984 NO. MONTHS	1985 % DAYS	1985 NO. MONTHS	1986 % DAYS	1986 NO. MONTHS	1987 % DAYS	1987 NO. MONTHS	1988 % DAYS	1988 NO. MONTHS	OVERALL PERCENT
181	14.8%	3	29.0%	6	0.0%	0	38.6%	5	63.8%	8	16.7%	2	832.1%
303	39.2%	7	47.7%	9	0.0%	0	48.2%	6	83.0%	10	0.0%	0	812.6%
305	60.8%	9	63.8%	9	0.0%	0	18.4%	3	68.5%	9	41.4%	5	791.4%
377	71.8%	12	35.9%	6	30.7%	4	99.7%	12	76.7%	10	0.0%	0	769.9%
479	55.9%	12	39.5%	6	85.5%	12	81.9%	11	69.0%	11	58.1%	7	736.4%
576	90.1%	12	64.4%	9	93.4%	12	99.7%	12	99.7%	12	58.1%	7	692.7%
577	93.2%	12	46.8%	8	80.5%	11	51.2%	9	50.1%	7	0.0%	0	635.0%
676	84.1%	12	53.2%	7	94.0%	12	99.7%	12	95.6%	12	58.1%	7	609.1%
677	88.5%	12	55.3%	7	94.0%	12	96.4%	12	100.0%	12	57.0%	7	553.2%
703	83.8%	12	35.1%	5	29.6%	6	49.9%	7	34.0%	5	12.6%	2	496.1%
707	65.5%	10	63.3%	9	95.3%	12	32.1%	4	84.7%	11	58.1%	7	479.0%
775	77.0%	12	58.1%	9	78.1%	12	34.8%	5	0.0%	0	26.3%	5	433.0%
802	81.9%	12	63.3%	9	0.0%	0	37.3%	6	74.5%	10	57.5%	7	412.4%
975	49.0%	12	0.0%	0	90.1%	12	95.1%	12	93.2%	12	9.0%	2	379.5%
976	92.6%	12	62.5%	9	69.6%	11	99.7%	12	75.1%	9	25.8%	4	344.6%

The accident frequency and accident rate series for passenger cars only are shown in Figures 6–9. The accident frequency and accident rate series for car-truck accidents are shown in Figures 10–13. Figures 14 and 15 show the fraction of car-truck accidents as a proportion of all accidents, which eliminates the need for an exposure base. Only frequencies are required for the calculation, yet it maintains a relationship to exposure by assuming that the changes in VMT will occur at the same relative rate for both cars and trucks.

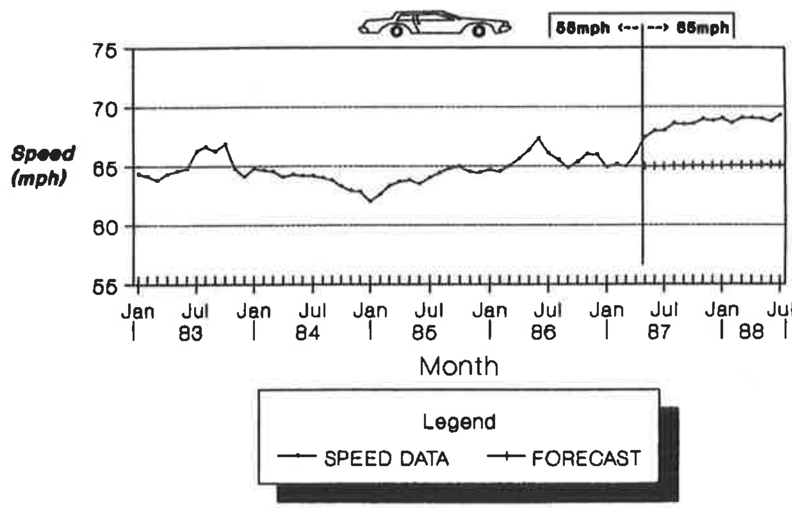


FIGURE 1 85th-percentile speed by month: passenger cars, all sites.

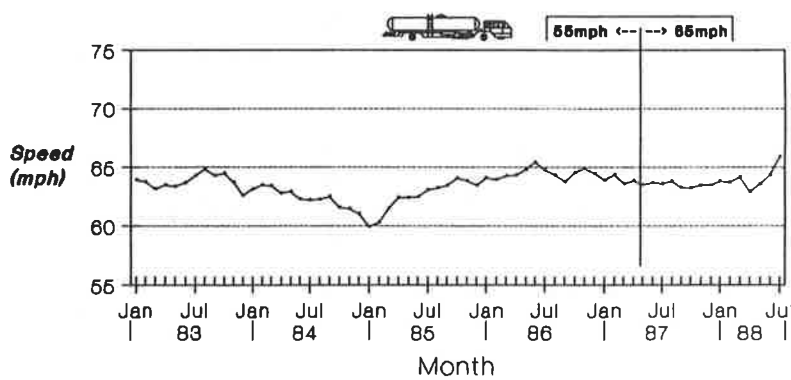


FIGURE 2 85th-percentile speed by month: trucks, all sites.

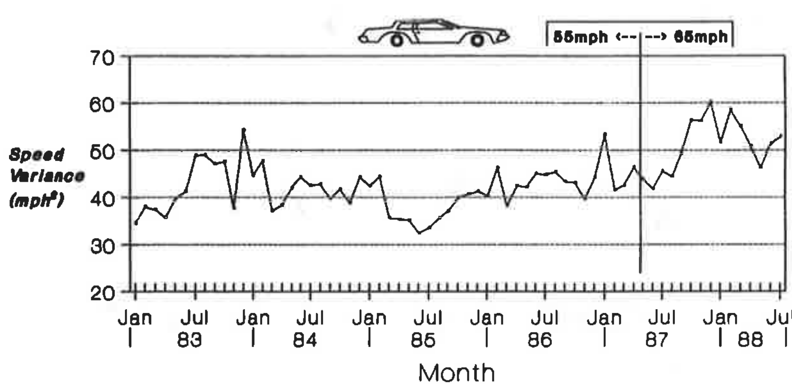


FIGURE 3 Speed variance by month: passenger cars, all sites.

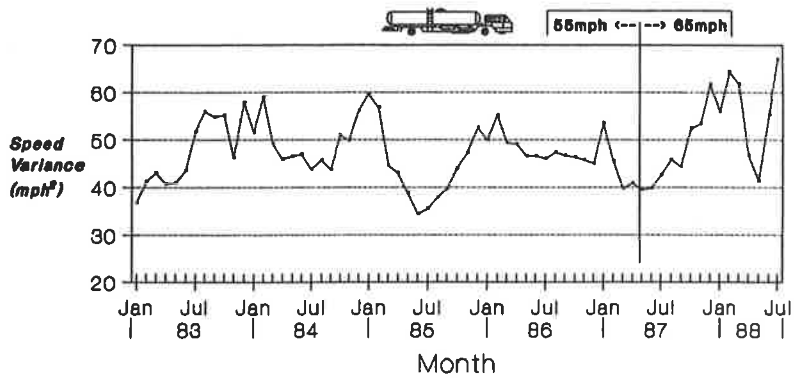


FIGURE 4 Speed variance by month: trucks, all sites.

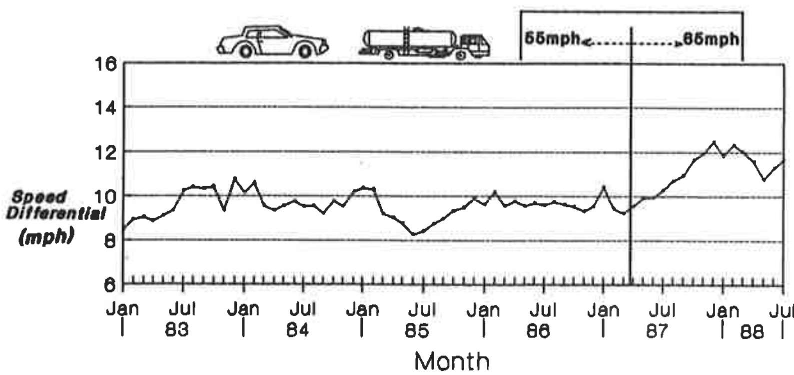


FIGURE 5 Car-truck speed differential by month: all sites.

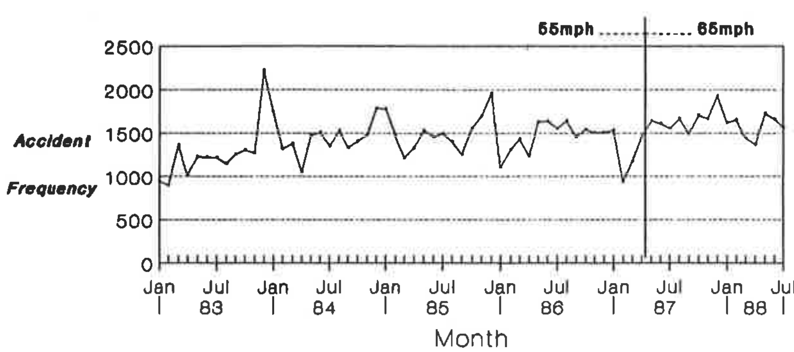


FIGURE 6 Accident frequency by month: all reported accidents, passenger cars, all sites.

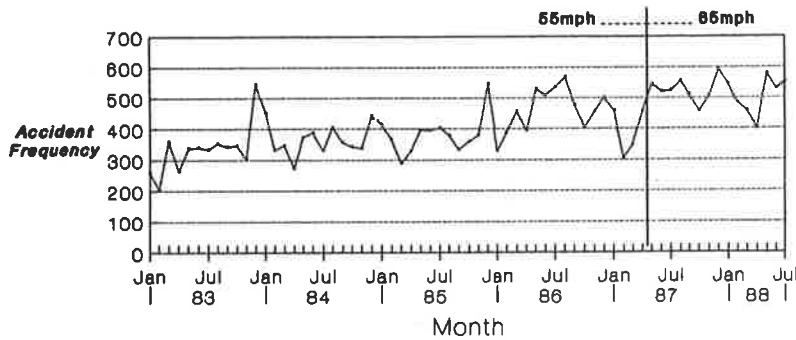
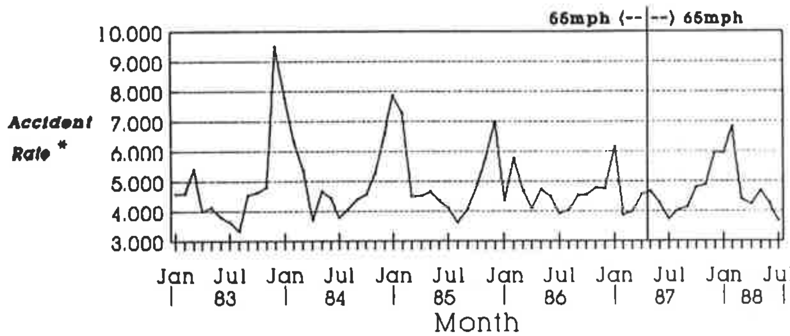
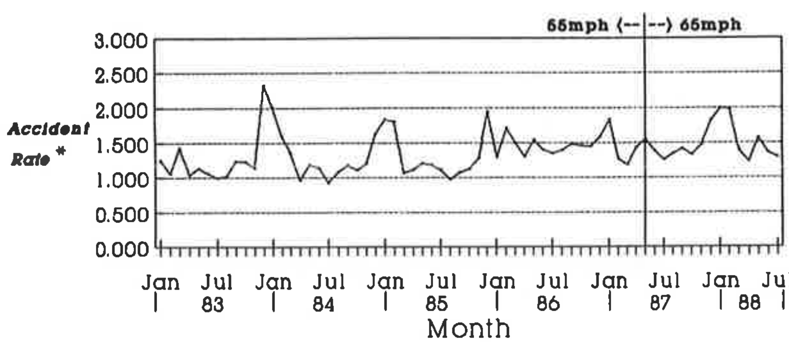


FIGURE 7 Accident frequency by month: fatal and injury accidents, passenger cars, all sites.



*
Accidents Per Month
Per Million VMT

FIGURE 8 Accident rate by month: all reported accidents, passenger cars, all sites.



*
Accidents Per Month
Per Million VMT

FIGURE 9 Accident rate by month: fatal and injury accidents, passenger cars, all sites.

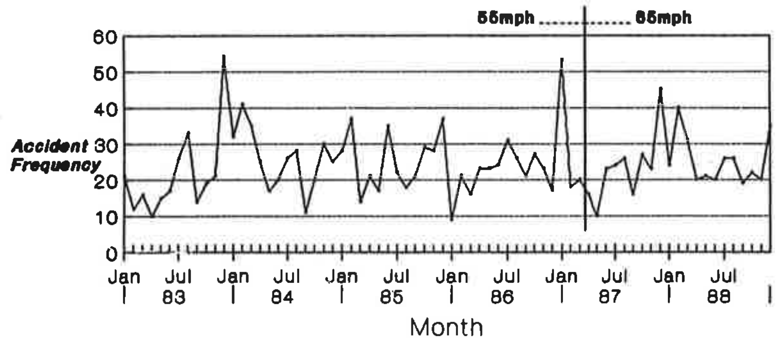


FIGURE 10 Car-truck accident frequency by month: all reported accidents, all sites.

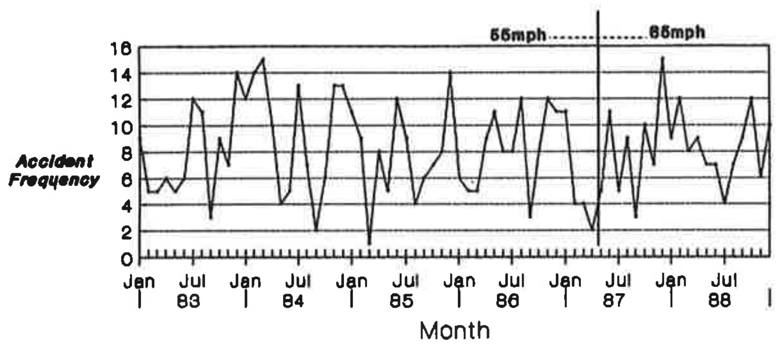
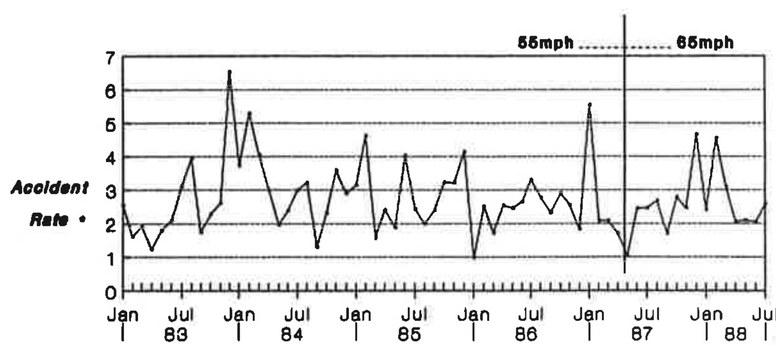


FIGURE 11 Car-truck accident frequency by month: fatal and injury accidents, all sites.



*
Accidents Per Month
Per Million VMT

FIGURE 12 Car-truck accident rate by month: all reported accidents, all sites.

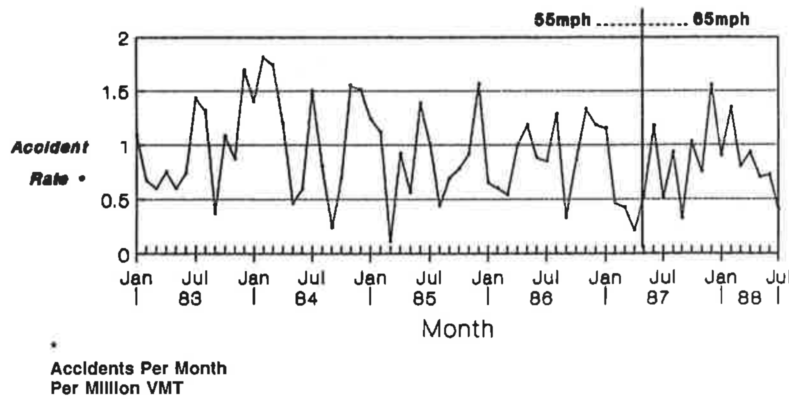


FIGURE 13 Car-truck accident rate by month: fatal and injury accidents, all sites.

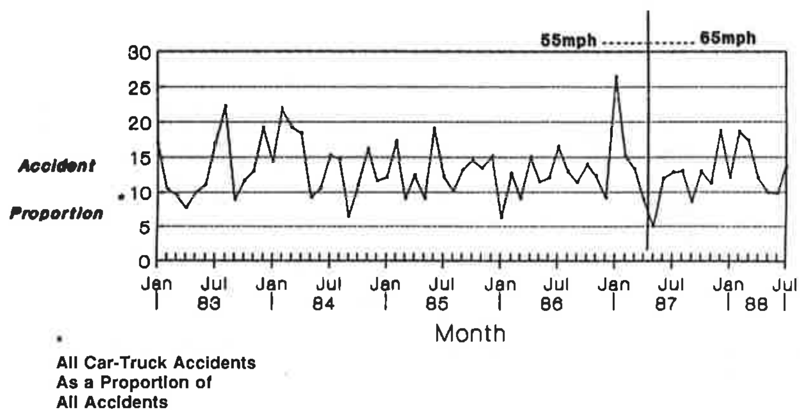


FIGURE 14 Car-truck accident proportion by month: all reported accidents, all sites.

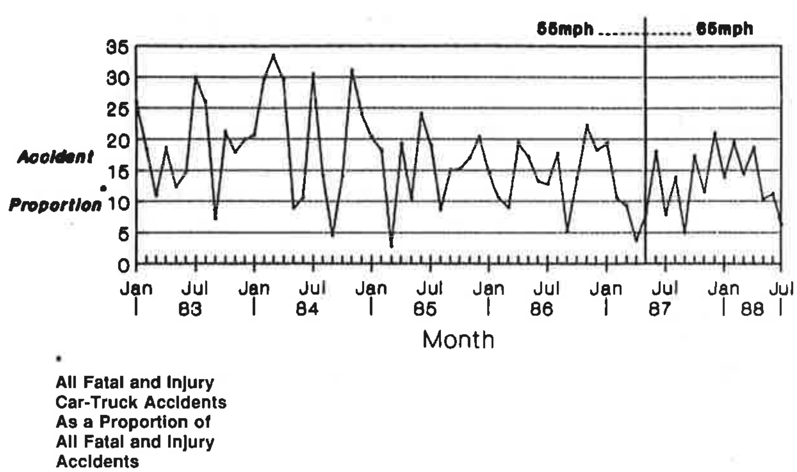


FIGURE 15 Car-truck accident proportion by month: fatal and injury accidents, all sites.

FINDINGS

Changes in Vehicular Speed

The change in the speed limit for rural Interstate facilities in Illinois occurred in May 1987. Treating the monthly observations of each measure as time series, ARIMA intervention analyses were performed to test for the presence, magnitude, and nature of changes, if any, in the speed data after the law change. The ARIMA analysis identified a statistically significant increase in the 85th-percentile speed for cars (see Table 3). The analysis also indicated that the form of the change was gradual and permanent, which suggests that the impact on the 85th-percentile speed occurred gradually after the law change. The ARIMA analysis also predicted that the ultimate increase in speed was 4.0 mph. No impact was detected on the speed variance of passenger cars. No permanent change in truck speeds was detected, measured either by 85th-percentile speed or by speed variance.

The speed differential between cars and trucks was also examined. The results suggested that a gradual, permanent increase occurred in the car-truck speed differential after the change in the law. The conclusion, however, is only supported at the 20 percent level of significance. The asymptotic value for the change was estimated at 2.8 mph.

Changes in Accidents

The accident rate time series covered the same 67-month period that was used for the analyses of the speed data (i.e., January 1983–July 1988). In addition to frequency and rate, a third accident measure for car-truck accidents was used: the proportion of car-truck accidents to all accidents.

The results of the ARIMA analysis of the accident measures, presented in Tables 4 and 5, were inconclusive. A statistically significant increase of 14.2 percent was found in the frequency of all accidents for the 65-mph rural Interstate sites. There was not, however, a corresponding increase in the frequency of fatal and injury accidents. The analyses did indicate an impact (at a low level of significance) on fatal and injury accident rates. A permanent increase in the fatal and injury

TABLE 3 IMPACT OF SPEED LIMIT CHANGE ON 85TH-PERCENTILE SPEED AND SPEED VARIANCE AT 15 SPEED-MONITORING SITES, CARS AND TRUCKS

	Speed Measures	
Vehicle Type	85th Percentile Speed (MPH)	Speed Variance (MPH^2)
Cars	delta MPH = +4.0[a] (GP, .05)[b]	No Impact
Trucks	No Impact	No Impact

[a] The delta value indicates the estimated asymptotic change in the 85th percentile speed.

[b] (GP, .05) indicates that a "gradual permanent" post-intervention model was identified at the .05 significance level. The reported significance level represented the least significant level among all relevant parameters in the ARIMA model used for this case.

TABLE 4 IMPACT OF SPEED LIMIT CHANGE ON ACCIDENT MEASURES FOR PASSENGER CAR ACCIDENTS ONLY, ALL SITES

	Accident Severity	
Accident Measure	All Accidents	Fatal and Injury Accidents Only
Accident Frequency (Accidents per Month)	Freq. Change = +14.2%[a] (AP, .05)[b]	No Impact
Accident Rate (Accidents per month) per Million VMT	No impact	Rate Change=18.5% (AP, .20)[a]

a The Freq. or Rate Change value for an "abrupt permanent" post-intervention model (AP) indicates the estimated permanent percent change in the accident frequency or accident rate that occurred at the time of the intervention.

b (XX, .xx) indicates the form and significance level of the post-intervention model; i.e.,

XX = AP (abrupt permanent), or .xx = .05, or
= GP (gradual permanent), or = .10, or
= AT (abrupt temporary). = .20.

Reported significance levels for each case represent the least significant level among all relevant model parameters for that case.

TABLE 5 IMPACT OF SPEED LIMIT CHANGE ON ACCIDENT MEASURES FOR CAR-TRUCK ACCIDENTS ONLY, ALL SITES

	Accident Severity	
Accident Measure	All Accidents	Fatal and Injury Accidents Only
Accident Frequency (Accidents per Month)	No Impact	No Impact
Accident Rate (Accidents per Month per Million VMT)	No Impact	Rate Change = -27.3%[a] (AP, .05)[a]
Accident Proportion[b] (Car-Truck Accidents as a % of All Accidents	No Impact	Prop. Change = -74.8%[c] (AT, .10)

a The Freq. or Rate Change value for an "abrupt permanent" post-intervention model (AP) indicates the estimated permanent percent change in the accident frequency or accident rate that occurred at the time of the intervention.

b Two proportions were examined. For all accidents, the proportion equaled all car-truck accidents divided by all accidents. For fatal and injury accidents only, the proportion equaled all fatal and injury car-truck accidents divided by all fatal and injury accidents.

c The Prop. Change value for an "abrupt temporary" post-intervention model (AT) indicates the estimated temporary percent change in the proportion of car-truck fatal and injury accidents among all fatal and injury accidents that occurred at the time of the intervention.

accident rate of about 18.5 percent was calculated (significant at the 20 percent level).

The time series analyses of car-truck accidents are presented in Table 5. Both the accident rate and the accident proportion measures, for fatal and injury accidents only, exhibited significant reductions in car-truck accidents associated with the law change.

CONCLUSIONS

On the basis of the analyses of speed data, there is evidence that an increase of 4 mph occurred in 85th-percentile speeds of passenger vehicles along rural Interstate highways in Illinois after the change to the 65-mph speed limit. There is not sufficient evidence, however, to indicate that any change occurred in truck speeds. There is also no strong evidence that either the speed variance or the car-truck speed differential changed.

The safety impact associated with the speed increase is not clear. Although there is strong evidence that the overall accident frequency increased more than 14 percent on these same facilities after the law change, it does not seem to have affected the injury and fatal accident experience. Furthermore, when traffic volume variations are taken into account, there is no strong evidence that there was an effect on accidents associated with the law change. Finally, there is no strong evidence that the car-truck accident experience was affected by the law change. In fact, the incidence of car-truck accidents decreased after the change.

The experience gained from this research suggests specific methodological issues for further study:

1. Sensitivity analysis of the study results on the particular month used (i.e., the increase in the speed limit);
2. Elimination of non-speed-related accidents; and
3. A study of long-term trends by extending the data collection period after the law change.

Most of the work being conducted today is global in its orientation. However, the variations from nationally aggregate study results found within this study, and initial indications that similar variations have been identified in other states, suggest the need for more analyses that are based on disaggregate data to allow the development of valid models for the preimplementation examination of proposed policies. Areas of research for further consideration include the following:

1. Comparison of speed and crash experience associated with the law change in states having similar roadway, environment, enforcement, and legal conditions;
2. Analysis of variations in the speed and crash experience

with respect to design attributes of the roadways, terrain, traffic mix, weather, and enforcement practices; and

3. A study of specific issues related to speed differential using headway and other data measured from pairs of vehicles, by type of vehicle.

ACKNOWLEDGMENTS

This paper is based in part on a project conducted by The Traffic Institute, Northwestern University, under an NHTSA grant. The authors acknowledge the support and assistance of Susan Partyka of NHTSA and David Skinner of the Transportation Systems Center. IDOT was helpful and cooperative in identifying and providing the data on which the analyses were based. The authors extend their thanks and gratitude to the many individuals at IDOT who were involved.

REFERENCES

1. *Special Report 204: 55: A Decade of Experience*. TRB, National Research Council, Washington, D.C., 1984.
2. A. F. Hoskin. Consequences of Raising the Speed Limit. *Journal of Safety Research*, Vol. 17, No. 4, Winter 1986, pp. 179–182.
3. R. Pfefer and W. Stenzel. *The Safety Impact of the 65 mph Speed Limit—A Time Series Analysis*. NHTSA, U.S. Department of Transportation, 1989.
4. G. E. P. Box and G. M. Jenkins. *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco, Calif., 1976.
5. R. McCleary and R. A. Hay. *Applied Time Series Analysis for the Social Sciences*. Sage Publications, Inc., Beverly Hills, Calif., 1980.
6. D. T. Campbell. From Description to Experimentation: Interpreting Trends as Quasi-Experiments. In *Problems of Measuring Change* (C. W. Harris, ed.), University of Wisconsin Press, Madison, 1963.

Publication of this paper sponsored by Committee on Traffic Law Enforcement.

Use of Speed Limiters in Cars for Increased Safety and a Better Environment

SVERKER ALMQVIST, CHRISTER HYDÉN, AND RALF RISSER

A major traffic safety problem exists in Europe because many road users do not comply with speed limits. Therefore, researchers at the Department of Traffic Planning and Engineering at Lund Institute of Technology in Sweden have developed a method for controlling speed at the source, that is, in the vehicle. This method involves use of a speed limiter (SL). At every change of existing speed limit, a sender emits impulses. Each vehicle is equipped with a receiver able to understand these impulses, automatically limiting the vehicle's maximum speed to the speed limit in question. This concept sounds utopic, and there is not the slightest chance of introducing an SL into road traffic unless the advantages, from safety and environmental viewpoints, clearly outweigh the possible disadvantages. Hypotheses about the SL's major impact on traffic were developed through a literature study, roundtable discussions, and self-observation studies. In addition to objective safety, these hypotheses concern security, travel time, road net capacity, energy consumption, air pollution, and noise. External test drivers are currently being chosen for behavior observations and interviews. Their attitudes concerning several safety aspects will be measured, then groups of drivers will use the SL for different lengths of time. After driving, they will be interviewed again about their attitude toward the SL and its practical use. Along with the field experiments, estimates will be made on aggregated levels. The results of these studies will be used in planning and designing a large-scale experiment that will be held as an option, unless there is a surprising change of status quo in the official attitude concerning an SL system.

In all countries of western Europe, a major traffic safety problem exists because car drivers do not comply with the speed limits set by authorities. This problem seems to be fairly independent of the type of limit (30, 50, 70, 90, 110, and 130 km/hr and those in between), and it must be viewed apart from theoretical discussions about which speed limits are adequate for different circumstances from a traffic safety perspective.

Besides simply putting up traffic signs, many methods to achieve better compliance with speed regulations have been tried, including new types of road design, humps and other hindrances, and law enforcement. These methods continue to be the ones most often used, along with arguments for better compliance with speed regulations through traffic safety campaigns. Some of these methods (e.g., speed humps) are efficient, but only locally so. At the same time they are considered to be too expensive to install in all the places they are needed. Some other methods (e.g., safety campaigns) are inefficient, or their efficiency is difficult to prove in the short term. Still another method—the use of radar pistols—is in opposition to a certain law in most European countries. This law disallows radar pistols because their use is considered to be of questionable validity in speed measurement and in the correct identification of the measured vehicle out of several vehicles.

Instead of discussing the problems with using these methods to adapt vehicle speed to existing limits, the problem has been approached from another angle. So far, little effort has been focused on the most natural solution, namely, to control the process at the source (i.e., in the vehicle). A method for accomplishing this goal is introduced in the following paragraphs.

THE SPEED LIMITER

The proposed method involves the use of a speed limiter (SL). At every change of existing speed limit, a sender emits impulses. Each vehicle is equipped with a receiver able to receive these impulses, automatically limiting the vehicle's maximum speed to the speed limit in question. This system would be automatic and obligatory.

Researchers in the Traffic Safety Group of the Department of Traffic Planning and Engineering at Lund Institute of Technology in Sweden have presented this idea to their sponsors, along with a plan for analyzing the feasibility of and the conditions for introducing SL equipment into road traffic.

It must be stressed that the focus at this stage is on built-up areas with speed limits of 30 to 70 km/hr. It is also important to stress that the system must be mandatory—that it eventually must be in operation all the time for all drivers. A guarantee that a significant majority of car drivers would be complying with speed rules would allow for a smooth introduction of the SL and would provide a real improvement. It is believed that voluntary use would not change speed behavior significantly.

The main problem in testing this method is obvious. Because such a system does not exist, it will be difficult and time-consuming to provide conditions that simulate an automatic and obligatory SL, even in limited areas. Thus, the studies presented in the following sections (i.e., studies already conducted and those planned for the near future) are theoretical and are based on answers to conditional questions

Department of Traffic Planning and Engineering, Lund Institute of Technology, University of Lund, P.O. Box 118, S-122 00 Lund, Sweden.

about options for future road traffic dealing with an entire speed-limiting system. Empirical studies are limited to detailed system aspects, such as technical reliability and functioning smoothness, as well as attitudes toward the use of speed limiters in cars.

PRECONDITIONS FOR INTRODUCTION OF AN SL

Demands to introduce an SL into road traffic will be acted on only if it can be clearly shown that the advantages, from a safety and environmental viewpoint, outweigh the disadvantages. Hence, an overall appraisal of the extent to which the SL can influence the traffic system is necessary. Following are some hypotheses about possible advantages and disadvantages of the SL.

Advantages

When drivers have difficulty judging their own speed and therefore drive faster than they intend, the SL would be of help. The speedometer does not always fulfill its function. Drivers frequently must look at other road users, the road, or other displays in the vehicle and cannot read the speedometer on these occasions.

When drivers are aware they are speeding but are doing so because they feel pressed to keep up with surrounding traffic, the SL should help them maintain the speed limit. This hypothesis is a typical one, and is of interest only in connection with the current traffic system. Obviously, if all vehicles were equipped with SLs this problem would be solved.

Sometimes, a driver's choice of speed is caused by irrational or emotional motives, for example, trying to be faster or stronger than other car drivers. In such cases, the SL can be assumed to be of help. However, the tendency to compete on the road may then find outlets other than the use of vehicle speed.

Overall, the SL will have a positive effect on safety, fuel consumption, air pollution, and noise pollution. If the SL leads to a more relaxed and smooth driving style, then the positive effects will be more pronounced.

Disadvantages

The SL limits the drivers ability to choose the vehicle's speed. This restriction in freedom of choice might lead to unwanted forms of behavior, at least in the beginning phase after introduction of the SL. As a psychological rule, compensating behavior almost always results when freedom of choice is reduced. Possible actions might include driving against red or more dominant behavior at pedestrian crossings.

Another interesting hypothesis is that the frustration that may result for drivers from a feeling of losing time would lead to even higher speeds in low-speed situations (e.g., right or left turns) where the maximum speed today is below what the SL would allow. Moreover, the possibility that problems might arise in connection with overtaking maneuvers must not be overlooked.

A significant increase in time consumption is another possible consequence of the use of an SL, assuming that drivers currently save some time by breaking certain rules.

COMPLETED STUDIES

Phase 1: Literature Study

The basic hypotheses about the SL's possible impact on car driver behavior have been guided by experts' interpretations of existing knowledge of car drivers' motives and actions (*1*). More detailed hypotheses were to be formulated in subsequent phases of the project.

Phase 2: Roundtable Discussions

A roundtable discussion is an open form of discussion that can be used to broaden knowledge of a subject not yet discussed widely. Agreements as well as contradictions among the participants are used to shed light on as many aspects of the matter as possible. Experiences with roundtable discussions (*2*) have indicated that the following advantages result from application of this method:

- Preset statements are avoided.
- Spontaneous comments are usual.
- New and independent trains of thought are more likely to be found.
- Goal setting is not based on a predetermined route choice.

Thus, increased opportunities arise for discovering aspects not previously considered.

Concerning psychological aspects, the roundtable discussions make clear that the question of acceptance, not explicitly included in the hypotheses named previously, has to be given priority (*3*). Moreover, it became more and more obvious that the aspects of technical reliability and functioning smoothness that should be considered thoroughly were precision, maneuverability, adjustability, comfort, and hamper resistance.

During the roundtable discussions, contacts were established with a German team of researchers also working on an SL project series (*4,5*). This group had built an SL that met the specifications for technical reliability and functioning smoothness. Of course, the SL must be manually operated at this time, and an automatic and compulsory system if far from being installed. However, an SL was purchased from Germany to allow simulation of the SL concept. In accordance with the research concept, it was then possible to start on the third phase of the project.

Phase 3: Self-Observation Studies

To begin, one car was equipped with the SL from Germany. Personnel employed at the Department of Traffic Planning and Engineering in Lund drove this car. During the trips, which they chose freely according to their needs, they used a built-in microphone to comment on unusual situations, as well as any change—positive or negative. All comments were immediately recorded on tape. In addition to these comments, expert discussions were held regularly at the department. The car was also equipped with a datalog for later recording of time, speed, and energy consumption. These studies were conducted to test the hypotheses resulting from Phases 1 and 2 to allow them to be better defined, completed, and put into operation.

The initial hypotheses are concerned with the individual drivers' acceptance of and reactions to the SL and with the effects on an aggregated level (e.g., safety, noise, and air pollution). As a main result of Phases 1–3, those hypotheses that deal with a situation in which only few cars are equipped were verbalized in more detail, as follows:

- Various degrees of resistance to the SL will disappear after an equipped car has been driven for some time (the length of time must still be established).
- Irrespective of the attitude shown before driving the equipped car, drivers will generally feel comfortable driving such a car. A systematic deterioration of attitude is not expected.
- There is a possibility that driving an SL-equipped car in today's traffic could be frustrating because the driver is pressed, or overtaken, so often.
- Drivers who want to drive slower, despite the pressure exerted by today's traffic, will find it much easier when their wishes are supported by the SL.
- On the other hand, the hurry reflected by today's traffic might influence drivers of SL-equipped cars, leading to such compensation tendencies as driving faster when turning left or right or driving against a red light.
- For drivers having a positive attitude toward lower speeds, the SL will be experienced as a support for their behavioral tendencies.
- For drivers having a negative attitude toward lower speeds, these attitudes will change for the better in the long run (as a dissonance phenomenon).
- A driver who is the only one using an SL in a fast-going traffic system might experience a kind of claustrophobia (because of being slower than others or not being able to overtake vehicles going slightly below the limit).
- No relevant safety gains for the system can be expected if only a few vehicles are SL-equipped. The safety gain for single drivers depends on their own attitudes, in response to the compensation tendencies. This effect should vanish after some time, however.
- Frustration with the SL could lead to revenge tendencies, because others cause pressure and possibly bad feelings at times. This effect should vanish after some time, as well.
- Compensation tendencies are not likely with individuals who have a positive attitude toward the SL.
- The reduction in fuel consumption and air pollution will be perceived immediately for people who have a positive attitude toward the SL. For those who initially have a negative attitude toward the SL, this advantage will be realized after some time.
- Time consumption will not be altered so significantly that usual routines will have to be changed on a large scale.
- The SL will cause smoother behavior because acceleration and deceleration will become less accentuated. The whole speed set will become slower, irrespective of attitudes toward the SL.

FUTURE STUDIES

In a previous section the reasons for the opinion that the SL must be mandatory were demonstrated. However, even if future work is basically directed toward effects within a fully equipped SL system, the transition phase to such a system might be long. Safety during this transition period might be influenced. The hypotheses just described, as well as short-term future work, must be viewed in this light.

Phase 4: Field Studies, Part 1

In Phase 4, a special test crew will be recruited. Crew members will drive the car, with and without the SL activated, simultaneously recording behavior. Additionally, time consumption, speeds, accelerations, and decelerations will be recorded automatically with help of the datalog that was calibrated in Phase 3. The driving period will be preceded by interviews to characterize the drivers' attitudes about driving, traffic safety, and traffic safety measures. The idea is to choose drivers with different attitudes toward traffic safety issues, especially toward the importance of speeds and measures to reduce speeds. After the test driving, new interviews will be conducted, focusing on reactions to the SL and changes in driver attitudes.

The tests of the hypotheses about driver behavior will be performed by specially selected drivers on a certain stretch of road in an urban area (e.g., in Malmö) and partly on the driver's usual route (e.g., the route driven to work). First, the drivers will drive these stretches a few times with the SL activated. When driving on the specially selected route, an observer will be in the car. The observer will take written notes on behavior, which cannot be covered automatically, and will ask questions about special situations that demand an explanation or interpretation of driver behavior. When driving the usual route, behavior will be recorded only by the datalog. The measurements will later be calibrated so that comparisons can be made between the before and after situations. Factors compared will include speed, acceleration, deceleration, and fuel consumption.

Primarily, these studies of behavior can be used to test hypotheses about drivers' reactions and behavior only when a few cars are equipped with SLs. The behavioral observations will be combined with detailed analyses of every recorded change in behavior and applied to whether this change will be influenced when an increased share of vehicles are equipped with SLs (including the scenario in which all cars are equipped). These analyses will be backed up with driver interviews after they have driven with an SL.

When the results of these studies give accurate enough estimates of driver reaction to the SL under given circumstances, work on the aggregated level will begin. The degree of accuracy in these estimates depends on several circumstances, of which the following are the most important:

- How well the test drivers represent the whole population,
- How well the chosen test routes represent the road network,

• Relevance of observed changes in behavior for situations in which the SL is mandatory and functions automatically, and
• Relevance of observed changes in behavior for situations in which many cars are equipped with SLs (including all being equipped and always in operation).

Of course, the accuracy of the estimates will be influenced by the size of the tests with hired drivers. However, the estimates will not become reliable until a large-scale, isolated experiment has been carried out.

Phases 5 and 6: Field Studies, Part 2; Surveys

Following are the most important hypotheses that are to be tested using the methods described:

• Normal attitudes toward speed-reducing measures are distinctly more sensible than politicians and traffic experts usually suggest.
• Among such groups as pedestrians, cyclists, the elderly, and beginning drivers, acceptance of the SL within a system frame will be high.
• In an SL system it will be easy to eliminate stress caused by the SL. Drivers will quickly get used to the SL because daily traffic will not remind them of former options (other drivers will not press and overtake).
• Stress will also be reduced because car drivers will stop trying to save time by driving fast. Such a concept has no basis in an SL system.
• Competition between drivers, and between drivers and other road users, will be reduced because one of the most important aspects of competition—namely, speed difference—is eliminated to a great extent.
• One important motive for sensationalist behavior is eliminated by eliminating speeding possibilities in an SL system. (It is hoped, however, that the optimistic hypotheses about compensatory behavior presented in the section on Phase 3 are correct.)
• Traffic in the frame of an SL system will become quieter and smoother. Fuel consumption, air pollution, and noise production will therefore decrease.
• Summing up all hypotheses about behavioral changes presented in the section on Phase 3, it is expected that safety in the SL system will be much higher than in today's system.

Phases 5 and 6 involve long-term work. Part 2 of the field studies will primarily concern a limited area or region in which a certain concentration of SL-equipped cars are in compulsory and automatic operation. Such a large-scale behavioral study will not be possible during the next 3 years, however, and the design of the study will depend on the results of Phase 4. If such a study can be carried out in 1994, some knowledge about a larger-scale functioning of an SL system will already be available. If all goes as planned, a study will be conducted in Germany by Von Winning and Krüger in which cars belonging to the federal state of Nordrhein-Westfalen and its counties will be equipped with an SL. The drivers of these cars will be instructed to use the SL in precisely defined areas. Although Von Winning and Krüger are not aiming at introducing a compulsory and automatic system in Germany, the study will come close to simulating a qualified number of car drivers in the region compulsorily using the SL in inhabited areas.

It is desired that a simulation study be conducted in Sweden. The knowledge gained from carrying out more behavior observations in different places and under different, but well-defined, conditions cannot be overestimated. Surveys, on the other hand, can be performed immediately after Phase 4, or at least earlier than large-scale field studies. It is difficult to draw valid conclusions about acceptance and other aspects of attitude from behavior observations alone. Even if those observations are done within a long-range experimental frame, they still have to be complemented with verbal data. Survey questions asked today must be of a hypothetical nature because not much is known about the characteristics of a future SL system. It is not possible to wait for results from large-scale experiments before asking questions. Sponsoring of this research will be probable only if hypotheses demonstrate that there will be sufficient acceptance of an SL system within a defined period of time by authorities, car manufacturers, and the public.

PERSPECTIVES

The most important perspectives of the prognoses will be traffic safety, travel time and traffic capacity, and environmental aspects (such as air pollution, energy consumption, and noise). Smooth functioning and technical reliability are not the subject of these prognoses because such aspects will not be problematic according to current knowledge.

Traffic Safety

Safety analyses are divided into two parts, namely, objective safety (reflected by accident risk and accident frequency) and subjective safety (perceived risks).

As long as there are no large-scale tests of an automatic and obligatory SL system, estimations of the final effects on accident risk can be done using only intermediate behavioral measures. Even this analysis may create problems. It is known, however, that changes in speeds usually are strong indicators of changes in accident risks. There are also some indications that red driving, accepted gaps, and similar behavior variables are related to accidents. Another behavior aspect that is of interest is the homogeneity in traffic, which primarily concerns any variance in speed between vehicles driving in the same direction (e.g., the frequency of overtaking maneuvers will be reduced decisively). A total homogeneity—that is, a flow without speed variance—would theoretically eliminate many accidents (especially rear-end and overtaking accidents). In general, there is likely to be a strong correlation between homogeneity and accident frequency. Also of interest is the study of how crossing pedestrians and cyclists choose gaps in the flow according to the homogeneity of vehicle speeds, as well as the safety implications of this factor.

Generally, it can be concluded that these behavioral measures are strongly linked to accident risk. Even though quantifications are partly missing, it should still be possible to make rough estimates of changes in accident risks by using behavioral measures.

The behavior types that reflect objective safety (accident risk) can be analyzed both from observations of test drivers' cars, which allow an assessment of the dangerousness of certain types of behavior (*6*) and from observations on the spot, in which traffic conflicts resulting from road-user interaction are used to indicate the existence and degree of accident risk (*7*). The former method can already be used when single cars are equipped with an SL (although system predictions must be made cautiously). The latter method will be adequate once large-scale experiments have started. In a somewhat longer perspective, it will also be possible to evaluate the total effects in a larger system with the help of accident analyses. (These analyses should, however, be supported by conflict and behavioral studies, partly to provide detailed information not possible to obtain from accident data for quite some time and partly to allow an explanation of the changes that occur.)

The self-observation studies have given some strong indications that the subjective safety of unprotected road users might increase. Drivers believed that their behavior toward pedestrians, for example, became more lenient. Such aspects can be observed to a certain extent, but questionnaire and review data will also be needed. The hired drivers should be asked about subjective safety aspects, as should drivers taking part in larger scale experiments or studies and, of course, pedestrians, cyclists, parents, residents, and so on. (Again, if asked today, many questions will be of strictly hypothetical nature.) Different groups of drivers might experience differences in safety as a consequence of the SL. For example, older and just-licensed drivers may react differently to an SL than would other drivers.

Travel Time and Traffic Capacity

Self-observation studies have initially shown that travel times are not much longer when an SL is used. It seems to be a matter of impatience if a driver wants to go faster, depending on whether the driver feels impeded by the SL. However, whether or not relevant losses of time could be a consequence of the introduction of an SL system is something that must be studied empirically. It is likely that the introductory studies of behavior, and the interviews focusing on acceptance and attitudes, will offer a good prediction of the way in which drivers will behave in traffic flows. This knowledge can then be applied to established models on the relationships between speed and gap choice and between travel time and capacity. It should then be fairly easy to predict what will happen on an aggregated level concerning these variables.

It is important to understand that these studies will not be sufficient in predicting the effects of a fully implemented SL system. It must be realized that interactions between road users in such a system might be influenced in quite a different way than when only a few cars are equipped. The effects cannot be appraised until large-scale experiments have been carried out.

Energy Consumption, Air Pollution, and Noise

Noise and air pollution are the environmental aspects that are of greatest interest. The effects of the SL on air quality can be studied when the first cars are equipped with SLs. The effects on an individual level can be studied either directly with the help of measurements of the test vehicles (e.g., an exhaust-measuring device connected to the datalog) or indirectly from knowledge of the relationship between speed and acceleration behavior and exhaust volumes. The effects on an aggregated level will be analyzed using the data from individual drivers as well as knowledge about aggregated speed data.

As explained previously, primarily the effects from few cars being equipped with SLs can be predicted. Earlier assumptions about changes of behavior following the transition from a system with few to one with many SL-equipped cars will be used. It will also be possible to make certain predictions about the effects of a fully implemented SL system. However, analogous to other aspects, it will not be possible to get reliable estimates until large-scale experiments have been carried out.

Energy consumption is closely linked to exhaust levels. The studies of energy consumption will, therefore, be planned in a similar way. Individual energy consumption will be measured directly in the test vehicles. Aggregations will be based on speed and acceleration data, including their distributions.

The noise aspect will be harder to study in the short range. It will be necessary to rely on theoretical models of the way in which noise levels are influenced by such factors as speed, acceleration, deceleration, and make of car. The possibility of aggregating is vague. A large-scale experiment is therefore important, partly because of an interesting aspect of construction in the SL that has already been acquired. This SL has an extra function that creates the possibility of decreasing the vehicle's acceleration capabilities as well as limiting the maximum speed. With the help of this function, the acceleration capabilities can be determined by choice. Thus, experiments can be performed with varying available acceleration capabilities. There is reason to believe that noise caused by powerful acceleration is common in urban areas. If so, limiting the acceleration capabilities will have a considerable effect on the noise level. This concept has yet to be examined.

CONCLUSIONS

Completed studies in this project as well as planned future research have been described. The presentation had to be kept short and unspecific in the current framework. Emphasis has been placed on describing the hypotheses and other aspects that are to be studied in the coming phases of the project, which methods could be used, and in which time perspective these phases should be done. The possibility of estimating a situation in which few cars were equipped with SLs and one in which all cars are equipped has been explored. As indicated, the need for continued research is great. It is therefore essential to proceed in a step-by-step fashion. In the first stage introspective studies were used to get a practical illumination of all aspects. The most important agenda was to revise, complete, or modify all behavioral hypotheses that had been formulated in the beginning of the project series. Changes in behavior are the ultimate indication of the way in which road users accept the SL and react to it and, thus, the consequences of the SL on safety, noise, air quality, travel time, and capacity.

After a first round of introspective studies with 10 to 15 employees of the Department of Traffic Planning and Engineering, external test drivers are now being chosen for behavior observations and interviews. Before driving, they will be asked about their attitudes toward traffic safety, traffic safety measures (especially speed-related measures), and the SL. On the basis of these interviews, the drivers will be grouped by their driving behavior and then compared. The behavior of the drivers when driving with and without the SL will be analyzed. Also, the motives behind the behavior will be considered to test the hypotheses presented in the section on Phase 3. At first, some of the drivers will drive the car for only a short time. Others will be driving for a longer time. This method will allow an analysis of how the adaption process works and whether the long-term effects are different from the short-term ones. This analysis will then be used in planning the rest of the field tests. After driving, the drivers will be interviewed again. In principle, the same questions will be asked, but some distinctions will be made when it comes to the drivers' attitudes toward the SL and the practical use of it.

Parallel with the field experiments, preparations will be made for making estimates on the aggregated level, that is, estimates for different phases within the transition phase and estimates in the fully implemented system. This work will begin with a theoretical study of models that can be used for estimating noise, air pollution, fuel consumption, capacity, travel time, and safety.

The results of these studies will be used to plan and design a large-scale experiment, which will be held as an option if evidence about the consequences of a hypothetical introduction of an SL system is not received by other means or unless there is a surprising change of status quo in the official attitude toward an SL system (e.g., in Sweden). However, it is essential that the study of attitudes toward traffic safety, safety measures, and the SL starts early. This research should include tests of hypotheses through the interviews of drivers of the experimental vehicles and through the questionnaire sent to a representative mix of people. The knowledge gained by these tests of hypotheses will be an important part of the base needed for decision making by authorities and industry.

To sum up, the next stage of research for the coming 3 years will include the following:

- Test driving with enough drivers in different situations to estimate the different effects well enough to enable large-scale testing as the next step;
- Estimates of the consequences for safety, noise, air pollution, and energy consumption using models, behavioral data, and interview data from the test driving; and
- Accumulating attitudes toward traffic safety issues, traffic safety measures in general, and the SL specifically.

The results from this stage will give a clear picture of the way in which individual drivers react to having a mandatory SL in the car. It is also considered possible to make a fairly good prediction of what will happen if the SL is introduced on a large scale, as well as actions needed to improve this knowledge.

The speed debate is intensive, and there is no indication that it will slacken in the near future. To date, the debate has focused primarily on rural roads. At the same time, the problems in built-up areas are of at least the same magnitude. There is also a lack of measures that can be expected to have a decisive influence on automobile speeds, especially in urban areas. The great potential for safety offered by the SL—according to the work done so far—means that the results from this next stage will be important for the future policies of authorities.

In the speed debate the performance of automobiles has been increasingly emphasized. Speed, acceleration, and handling abilities of modern cars, in combination with a quiet and comfortable inner environment, naturally influence the choice of speed. In this perspective it is important that car manufacturers also get as accurate a picture as possible of the qualities of the SL, how it will be accepted, and what consequences it will have for behavior, safety, noise, and energy consumption. The manufacturers can thereby adjust the production to better meet new demands that authorities, as well as consumers, might specify.

REFERENCES

1. C. Hydén and S. Almqvist. Förarbestämt förhandsval av fordonets maximihastighet. *Litteraturöversikt och problemanalys*, Institutionen för Trafikteknik, LTH, Lund, Sweden, 1987.
2. T. Bauer, R. Risser, P. Soche, W. Teske, and C. Vaughan. Kommunikation im Strassenverkehr. In *Zeitschrift für Verkehrsrecht* 26, Heft 6, Vienna, Austria, 1981.
3. S. Almqvist, C. Hydén, and R. Risser. *A Speed Limiter in a Car: Phase 2, Hypotheses Concerning Effects on Driver Behavior and Interaction.* Lund Institute of Technology, Lund, Sweden, 1991.
4. Von Winning and M. Krüger. *City-Paket und Geschwindigkeitsschalter. Verkehrsberuhigung am Auto.* ILS Schriften Nr. 35, ILS Dortmund, Germany, 1989.
5. S. Picot. *City-Paket und Geschwindigkeitsschalter, Band C: Fragen der Akzeptanz—Psychologische Aspekte.* Institut für Stadtplanung und Teknikberatung, Munich, Germany, 1989.
6. R. Risser. Behavior in Traffic Conflict Situations. *Accident Analysis and Prevention*, Vol. 17, No. 2, 1985.
7. C. Hydén. *The Development of a Method for Traffic Safety Evaluation: The Swedish Traffic Conflicts Technique.* Bulletin 70. Institutionen för Trafikteknik, LTH, Lund, Sweden, 1987.

Publication of this paper sponsored by Committee on Traffic Law Enforcement.

Analysis of Vehicle Operations on Horizontal Curves

DONALD W. REINFURT, CHARLES V. ZEGEER, BRENT J. SHELTON, AND TIMOTHY R. NEUMAN

A study was conducted to quantify the effects of horizontal curve features on such operational variables as changes in vehicle speeds and vehicle encroachments over the centerline and edgeline. This information was considered important in determining curve design criteria that would lead to effective safety and operational improvements at current curve sites. The data included geometric, traffic, and operational measures from a data base of 78 curve sites in New York State. Various statistical procedures were used, including linear regression analyses, analyses of operations by differing groups of geometric conditions, residual analyses, and locally weighted nonparametric regression. It was found that, as curves become sharper, there is a proportionally greater increase in speed reduction and edgeline encroachments on the inside lane (i.e., the lane on a curve where the motorist must steer to the right). Centerline encroachments in the outside lane increase more drastically than those on the sharper curves to the right. These findings support the contention of drivers cutting the curve short, which can result in run-off-road crashes on the inside of the curve as well as head-on and opposite-direction–sideswipe crashes with oncoming motorists. Appropriate curve design guidelines are discussed that may help to minimize these operational and potentially accident-related trends.

One of the major safety and operational problems facing motorists on two-lane rural highways relates to horizontal curves. Studies have consistently found that curves account for a higher rate of crashes and for greater crash severity than tangent sections of highway (*1,2*). The increased crash rates on curves are to be expected, because a curve requires a driver to perceive a change in roadway alignment and to take appropriate action, such as braking and steering changes. On sharp curves or under adverse environmental conditions (e.g., at night during rain or in fog), these tasks can be quite difficult.

Highway agencies are responsible for identifying curves that have safety and operational problems and for making necessary improvements. Although crash data may be useful in providing insights into such problems, crashes are typically rare at a given curve site. In fact, Zegeer et al. (*1*) found an average of only about one crash per curve at 10,500 curve sites on rural two-lane roads in Washington State during a 5-year period. Thus, the sample of crashes at a given site is often small and most likely will not reveal the full picture of safety and operational problems for drivers who encounter the curve.

The collection and analysis of traffic operational data on a sample of curves can be useful in several respects. For example, it can provide insights into how drivers react to various curve geometrics. Traffic operational measures on curves can also indicate the adequacy of the curve design in handling the traffic mix on the curve and possibly help to suggest accident problems that may result. Such knowledge can be useful in selecting appropriate roadway improvements on existing curves and better designs for alignment on new roadways (*3*).

The operational characteristics of various horizontal curve features were investigated. The operational measures of concern included speed changes and vehicle encroachments over the centerline and edgeline. Figure 1 shows edgeline and centerline encroachments in a curve to the right (i.e., the inside of the curve). Geometric and roadway features that were analyzed included degree of curve, length of curve, superelevation deviation, vertical alignment (grade), and roadside hazard. The data were taken from a data base of 78 curve sites in New York State and included geometric, traffic, accident, and operational variables.This data base was developed by Terhune and Parker (*4*). All 78 study sites were on two-lane rural roadways. Various statistical procedures were used to correlate operational measures with curve geometrics, resulting in recommendations for improved curve designs to minimize the operational problems that were identified.

BACKGROUND

Numerous studies have been conducted in recent years to record vehicle operations on curves. Glennon et al. (*2*) monitored vehicle speeds and lateral placement through five horizontal curves in Illinois and Ohio. They found that some drivers overshoot the curve radius, producing minimum vehicle path radii sharper than the highway curve. This tendency was found to be independent of vehicle speed.

Vehicle speed data were also observed by Glennon et al. for free-moving vehicles as they traversed 60 curve approaches. The sharpness of the impending curve was the factor most associated with speed changes by the drivers. Drivers tended to begin adjusting their speeds only as the curve became imminent, and speed reduction increased linearly with increasing degree of curve. Only a slight difference in speed changes was found for narrow versus wide roadways.

D. W. Reinfurt and C. V. Zegeer, Highway Safety Research Center, and B. J. Shelton, Department of Biostatistics, University of North Carolina, Chapel Hill, N.C. 27599. T. R. Neuman, Jack E. Leisch Transportation Group, CH2M Hill, 1890 Maple Avenue, Evanston, Ill. 60201.

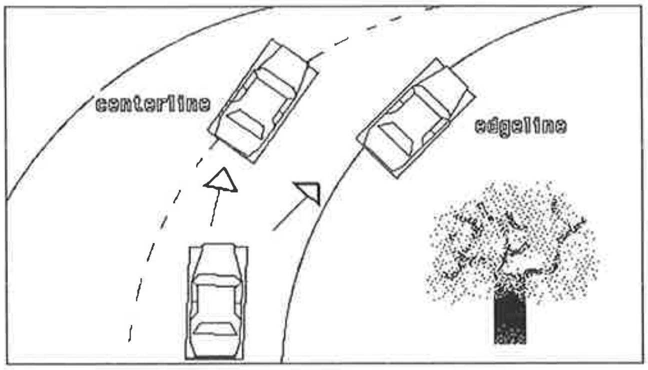


FIGURE 1 Centerline and edgeline encroachments in a curve to the right.

Jennings and Demetsky (*3*) collected traffic volume, vehicle speed, and lateral placement data at five curve sites in Virginia. Although their primary objective was to evaluate the effects of post-mounted delineator systems, they also analyzed driver responses in general at eight other sites. Similar average speeds, vehicle placements, and centerline encroachments were observed at various sites that had similar delineators. Vehicles were found to travel further from the roadway edge when delineation was present.

Terhune and Parker (*4*) collected various operational measures at 78 horizontal curve sites. Those measures found to be related to accident rate included traffic volume, average speed reduction, rate of centerline encroachments, and edgeline encroachments. However, none of the operational measures were incluided in the best-fitting accident predictive models.

Datta et al. (*5*) collected several operational measures at 25 rural, two-lane curve sites as part of a study on accident surrogates. Speed differential and traffic volume were the only operational measures included in any of the best-fitting accident predictive models, along with several nonoperational variables (e.g., sideslope, degree of curve, and superelevation error).

Several studies have also used operational measures to evaluate the effectiveness of curve delineation treatments. For example, Rockwell et al. (*6*) evaluated curve modifications on horizontal curves. Operational measures collected for traffic stream drivers included vehicle speed profiles and severe lateral displacement. Although signing treatments were largely ineffective, reductions in speed variance resulted from some of the modifications, such as pavement markings.

Rockwell and Hungerford (*7*) tested six delineation treatments at rural curve sites in Ohio. Some treatments (e.g., post delineators, raised pavement markers, and transverse pavement stripes) were recommended at selected curve sites with a high rate of nighttime crashes and a high proportion of transient drivers.

Although a number of studies have been conducted on vehicle operations at horizontal curve sites, few have analyzed the influence of various curve design features on vehicle operations. These geometric features of curves and their effects on vehicle operations are addressed in the following paragraphs. The results of the analysis are then used to recommend design improvements for new and existing horizontal curves.

DATA

As explained previously, the data base was developed by Terhune and Parker for a study to identify accident surrogates, that is, to find which geometric and operational measures were related to accident experience and therefore could be used as measures of high accident potential at curve sites. In that study, accident prediction models were developed with various geometric and operational measures as independent variables, resulting in degree of curve and average daily traffic (ADT) being the variables most strongly related to crash rates. In the study reported here, accident data were not used, but relationships between roadway and geometric features and operational measures were determined.

The field operational data were typically collected during daylight hours on dry pavement. Encroachment and speed change data were collected by a two-person crew from a concealed location at each site. Operational (i.e., dynamic) data were collected in half-hour increments until 4 hr of data were collected per curve site. Traffic volumes and edgeline and centerline encroachments were counted for each lane. An encroachment was recorded each time a vehicle tire passed over the roadway edgeline or centerline on the curve. Encroachment counts were divided by traffic volume to yield encroachment rates. Because encroachments obviously occur from either the inside lane or outside lane of a curve, the encroachment data were categorized according to these lane characteristics.

Speed data were also collected on selected isolated vehicles—those with a headway of 9 sec or more between leading and trailing vehicles—using a radar meter at points 250 ft before the midpoint of the curve and at the midpoint. The desired measure of speed reduction was the difference between the two measurements.

As mentioned, the geometric and roadway features that were analyzed (referred to as nonoperational variables) included degree of curvature, length of curve, superelevation deviation, vertical alignment (grade) and roadside hazard rating. The roadside hazard rating was a 6-point rating scale, which described the roadside conditions within 15 ft of the edge of the paved surface. The scale was defined as follows:

Roadside Condition	*Rating*
Clear, no fixed objects, fairly level terrain	1
Vegetation or yielding objects, no rigid fixed objects, fairly level terrain	2
Isolated rigid fixed objects, fairly level terrain	3
Ditch throughout most of curve, no embankment or sideslope > 3:1	4
Embankment or sideslope > 3:1	5
Numerous or continuous rigid fixed objects	6

The operational measures used were selected because they are important in determining how drivers maneuver their vehicles on curves of differing roadway and geometric conditions. Such measures are not necessarily related directly to vehicle crashes, although there are at least some logical associations between unsafe vehicle maneuvers and the likelihood of a collision. For example, frequent or several encroachments over the centerline could lead to head-on crashes, particularly on curve sites with moderate-to-high volumes of oncoming vehicles. High rates of edgeline encroachments may indicate that drivers are approaching the curve too fast or are surprised by the sharpness of the curve. Speed change is related to both the approach speed and the speed in the curve. Large speed reductions by vehicles going into a curve indicate that drivers are decelerating quickly to negotiate the curve. Such a driving task may be even more difficult at night or under wet or icy pavement conditions, which could indicate a high potential for crashes.

ANALYSIS AND RESULTS

Descriptive Analyses

Descriptive analyses were conducted for the New York State surrogate data base. Frequency listings for operational and nonoperational variables of interest were examined, as well as Pearson chi-square (X^2) measures of association. The frequency listings proved useful in determining the formation of category levels for certain variables later categorized. Pearson X^2 measures of association were calculated between each operational variable and each nonoperational variable, and the significances of each association was determined (see Table 1). Included in the table are both the value of X^2 divided by the degrees of freedom (to normalize values for comparisons among cells in the table) and the p value for examining the strength of the association.

Average speed reduction was found to be significantly associated with degree of curvature (for $\alpha = 0.05$). Centerline encroachments for curves to the right were significantly associated with grade. Centerline encroachments for curves to the left were significantly associated with degree of curvature, curve length, and grade. Edgeline encroachments on the inside lane were significantly associated with superelevation deviation and marginally significantly associated with degree of curve ($p = .095$) and grade ($p = .062$). Edgeline encroachments on the outside lane had no significant association with any of the nonoperational variables.

TABLE 1 MEASURES OF ASSOCIATION BETWEEN OPERATIONAL AND NONOPERATIONAL VARIABLES

Roadway Feature	Speed Reduction	Centerline Encroachment		Edgeline Encroachment	
		Inside	Outside	Inside	Outside
Degree of Curve	3.6* 0.03 +	2.1 0.33	6.1 0.002	2.4 0.095	0.7 0.50
Curve Length	0.3 0.72	0.5 0.61	4.5 0.01	1.5 0.22	0.6 0.55
Superelevation Error	2.7 0.07	1.1 0.34	0.5 0.63	6.2 0.04	0.5 0.58
Shoulder Width	1.8 0.16	0.7 0.49	0.7 0.50	0.7 0.52	0.6 0.55
Grade	1.1 0.30	10.8 0.001	6.3 0.01	1.7 0.06	0.0 0.92
Roadside Hazard: Outside	1.1 0.32	1.2 0.30	0.7 0.52	0.3 0.73	0.1 0.88
Roadside Hazard: Inside	2.3 0.10	2.4 0.10	0.3 0.74	0.5 0.58	0.3 0.74

$*X^2/df$ +p-value

As an alternative method for exploring the data, averages for each of the nonoperational variables were compared across low and high rates of the operational (or outcome) variables. After the values of the four encroachment variables were normalized by traffic volume (number of vehicles per hour passing through the curve in either the inside or outside lane yielding encroachment rates), they along with speed reduction were dichotomized across each variable. The lower values (values below and including the median value of each variable) comprised one group, whereas the other group contained the higher values (values above the median). Table 2 presents means and standard errors for each of the seven nonoperational variables within the two subgroups of each of the operational variables. This breakdown allowed for visual comparison of the means of the nonoperational variables between high and low categories for each operational variable. In a sense, this method is similar to discriminant analysis in that it examines levels (averages) of the nonoperational variables that are associated with low versus high outcome (operational) groups.

In Table 2, consider the operational measure of average speed reduction. Of the sample curve sites, half had low speed reductions (below 1.7 mph) and half had high speed reductions (at least 1.7 mph). Average curvature, curve length, and other roadway features were computed within these low and high categories of speed reduction. For example, curve sites with low speed reductions had an average curvature of 4.22 degrees, compared with 6.42 degrees for the curve group with high speed reductions. In other words, greater speed reductions were associated with sharper degrees of curve.

In Table 2, the greater speed reduction group is associated with shorter average curve length (714.9 ft) than the lower speed reduction group (764.7 ft), probably because greater speed reductions occur with sharper curves and because sharp curves are usually shorter than mild curves. Other roadway features associated with greater speed reductions include greater superelevation deviation (i.e., more superelevation deficiency), wider shoulders, and steeper grades. The effect of roadside hazard on speed reduction is unclear.

Similar types of data summaries are presented in Table 2 for centerline and edgeline encroachments. Higher centerline encroachment rates for curves to the right were observed for narrower shoulders and steeper grades. Also, curves with higher roadside hazard ratings resulted in greater centerline encroachment rates for curves to the right, perhaps because motorists tend to shy away from the edge of the pavement on curves where roadside hazards are greater (e.g., large trees or steep sideslopes adjacent to the roadway).

Centerline encnroachment rates for curves to the left are also presented in Table 2. Higher encroachments occur for sharper curves, shorter curves, narrower shoulders, and steeper grades. These results indicate that vehicles in curves to the left or right are more likely to encroach the centerline when confronted with curves having more restrictive geometrics.

The frequent occurrence of centerline encroachments on sharp curves deserves further discussion. Crash data reveal that run-off-road crashes are much more frequent on horizontal curves than are head-on crashes. For example, Zegeer et al. (*1*) found that fixed-object plus rollover crashes accounted for 57.1 percent compared with only 8.2 percent for head-on plus opposite-direction–sideswipe crashes. Thus, a

TABLE 2 MEANS AND STANDARD ERRORS OF NONOPERATIONAL VARIABLES DICHOTOMIZED BY OPERATIONAL VARIABLES

Roadway Feature	Average Speed Reduction: Curve to the Left (mi/h)		Centerline Encroachment Rates (No/Hr/ADT)				Edgeline Encroachment Rates (No/Hr/ADT)			
			Curve to the Right		Curve to the Left		Curve to the Right		Curve to the Left	
	<1.7* (31)**	1.7+ (31)	<0.13 (27)	0.13+ (26)	<0.24 (31)	0.24+ (31)	<0.24 (32)	0.24+ (27)	<0.11 (37)	0.11+ (26)
Degree of Curve	4.22*** 0.29+	6.42 0.60	4.69 0.35	4.68 0.36	4.35 0.32	6.13 0.60	4.09 0.23	5.29 0.42	5.38 0.52	5.09 0.45
Curve Length	764.7 63.52	714.9 52.28	788.1 57.86	758.5 69.76	875.0 56.67	639.0 58.38	844.4 59.99	700.0 64.72	719.0 46.15	806.5 81.24
Superelevation Error	0.04 0.004	0.05 0.003	0.04 0.004	0.04 0.004	0.04 0.004	0.04 0.003	0.05 0.004	0.04 0.003	0.04 0.003	0.04 0.004
Shoulder Width	7.74 0.29	8.48 0.31	8.56 0.25	7.85 0.36	8.39 0.23	7.88 0.36	8.41 0.29	8.00 0.34	8.11 0.24	8.15 0.39
Grade	1.39 0.11	1.65 0.14	1.26 0.11	1.62 0.15	1.35 0.12	1.66 0.13	1.19 0.11	1.69 0.14	1.38 0.11	1.69 0.15
Roadside Hazard Outside	3.81 0.29	3.45 0.33	3.26 0.34	3.80 0.36	3.52 0.35	3.66 0.27	3.26 0.33	3.80 0.37	3.84 0.30	3.23 0.31
Roadside Hazard Inside	3.48 0.28	3.65 0.36	3.07 0.33	3.65 0.36	3.71 0.32	3.41 0.32	3.37 0.35	3.35 0.35	3.68 0.31	3.38 0.32

*50 percent of the sample had speed reductions of less than 1.7 mph

**Number of curves in operational variable group

***Mean +Standard error Data not available for several sites

greater number of edgeline encroachments (i.e., related to run-off-road crashes) than centerline encroachments should be expected.

To help explain this apparent inconsistency, it should be considered that many of the centerline encroachments may be controlled, or intentional, encroachments. In other words, on a curve to the left, some drivers will intentionally cut the corner while driving across the centerline if they see no oncoming vehicles. Although this maneuver seems dangerous, many rural roads have low traffic volumes and thus opposing vehicles are relatively infrequent. Further, if a vehicle encroaches the centerline in a curve to the right because of excessive speed while another vehicle is approaching, the driver of the oncoming vehicle can often take evasive action and avoid the crash.

The factors related to edgeline encroachments for vehicles on curves to the right and left are also presented in Table 2. Factors related to higher edgeline encroachment rates for curves to the right include sharper curves (5.29 versus 4.09 degrees), shorter curves (700 versus 844 ft), slightly narrower shoulders (8.0 versus 8.4 ft), and steeper grades (1.69 versus 1.19 percent). For vehicles on curves to the left, higher edgeline encroachments were found for steeper grades (1.69 versus 1.38 percent). However, sharper curves were surprisingly not associated with more edgeline encroachments. As mentioned previously, for vehicles on curves to the left, sharper curves (6.10 versus 4.35 degrees) were associated with greater centerline encroachments, which would be consistent with the lack of increased edgeline encroachments for those vehicles. This finding may seem counterintuitive because vehicles in sharper curves to the left may be expected to run off the road on the right rather than encroach the centerline. Although such edgeline encroachments occur with some frequency (and sometimes lead to run-off-road crashes), the drivers appear to cut the curve more often than encroach the edgeline for these sharper curves. This finding may also indicate that, although edgeline encroachments are relatively rare, their occurrence may more often result in a crash than centerline encroachments because the latter will lead to a multivehicle crash only if an oncoming vehicle is present.

The relationship of edgeline encroachments to crashes may depend largely on roadway width. An edgeline encroachment on a curve with narrow lanes (e.g., no more than 10 ft) and no shoulder can often result in a driver's loss of control and a fixed-object crash, particularly if the sideslope is steep or there are fixed objects near the roadway. On curves with wide lanes and wide paved shoulders, an edgeline encroachment presents a much lower chance of a crash.

The key results of Table 2 are summarized in Table 3. The influence of various roadway features on vehicle operations can easily be seen. For example, as discussed previously, the greater amount of average speed reduction is associated with sharper and shorter curves, more superelevation deviation, and steeper grades. Likewise, steeper grades on curves are associated with an increased incidence of each of the operational measures, which is important in explaining operational problems on curves. An increased level of hazard in curves to the right is related to an increased incidence of vehicles on these curves that encroach the centerline. This finding may indicate a tendency to steer away from trees, steep slopes, and other roadside hazards on the right when going around a curve to the right.

TABLE 3 SUMMARY OF ROADWAY FEATURES ASSOCIATED WITH HIGHER INCIDENCE OF OPERATIONAL MEASURES

Roadway Feature	Average Speed Reduction	Centerline Encroachments		Edgeline Encroachments	
		Curve to The Right	Curve to The Left	Curve to The Right	Curve to The Left
Sharper Degree of Curve	•		•	•	
Shorter Curve Length	•	•	•	•	
More Superelevation Deviation	•				
Narrow Shoulder Width		•	•	•	
Steeper Grade	•	•	•	•	•
Greater Roadside Hazard: (Outside Lane)		•		•	
Greater Roadside Hazard: (Inside Lane)		•			

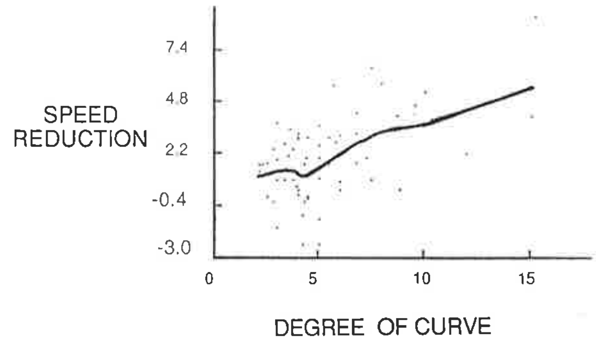


FIGURE 2 Speed reduction versus degree of curve.

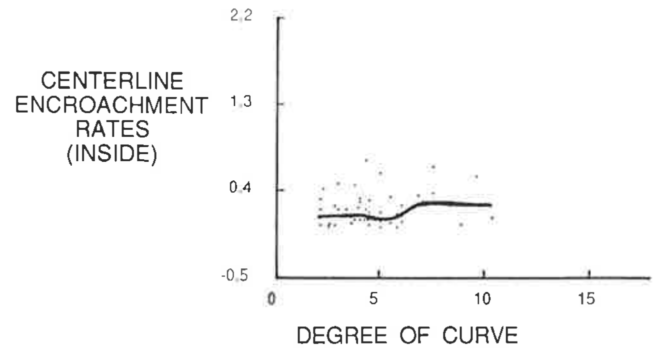


FIGURE 3 Centerline encroachment rates on curves to the right versus degree of curve.

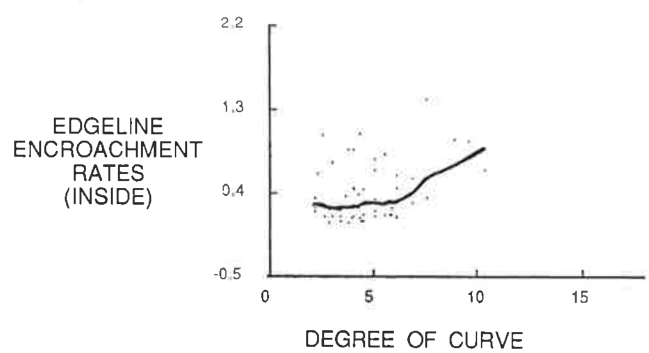


FIGURE 4 Edgeline encroachment rates on curves to the right versus degree of curve.

Regression Analyses

Analysis for Entire Sample of Curves

Most of the measures of association discussed previously are essentially correlation statistics, which are useful indicators in determining which nonoperational variables (e.g., degree of curve) might best account for the amount of variation in the operational variable values (e.g., speed reduction). To help explain this variation, least squares regression analyses were carried out using the information obtained from the Pearson X^2 analyses.

Scatterplots of actual values were constructed of each operational variable (vertical axis) versus each promising nonoperational variable (horizontal axis). Scatterplots using degree of curve as the nonoperational variable are shown in Figures 2–4 for speed reduction, centerline encroachments for curves to the right, and edgeline encroachments for curves to the right, respectively. Included in each figure is a nonparametric regression curve fitted by a locally weighted regression procedure (LOWESS), which is discussed later.

A variety of models were considered, but only two are noteworthy. On the basis of R^2 values and significance of parameter estimates (using SAS PROC GLM), these two models were centerline encroachments for curves to the left versus degree of curve and average speed reduction versus degree of curve—with R^2 values of 0.37 and 0.30, respectively. Because relationships are being investigated rather than using regression as a prediction model, the modest R^2 values are acceptable.

Analysis Within Mild Versus Extreme Roadway Conditions

It was hypothesized that a more definitive relationship might be seen between the operational and nonoperational variables if the curves were dichotomized into seemingly mild versus hazardous conditions. Thus, the data were categorized into two groups of curves on the basis of certain ranges of the following nonoperational variables:

- Superelevation deviation [termed "superelevation error" by Terhune and Parker (*4*)],
- Grade, and
- Roadside hazard rating.

The two groups created from these three variables are referred to as being favorable or unfavorable. The unfavorable group consisted of curves meeting one or more the following criteria:

- Superelevation error greater than 0.05,
- Grade rating of 3 (i.e., very steep), or
- Roadside hazard rating of 6 (i.e., most hazardous).

Plots were constructed within each group using degree of curve as the main nonoperational variable of interest (see Figures 5 and 6). As expected, associations between the degree of curve and the operational variables were much stronger in the unfavorable group. However, the data are still rather dispersed, pointing to the need for an analysis of residuals (i.e., differences between observed and predicted values).

Residual Analyses

Residual analyses were conducted for the univariable models formulated in the favorable and unfavorable curve groups. These analyses attempted to determine whether certain least squares regression assumptions were invalidated. Plots of the residuals versus the observed values for the combined curve group suggested that there is a violation of one of the regression assumptions because the error variable increases with increasing values of degree of curve. This tendency is indicated by the shape of the residual plot in Figure 7, which, in turn, suggested the use of a variance stabilizing transform. However, these transforms failed to improve on the least squares regression models.

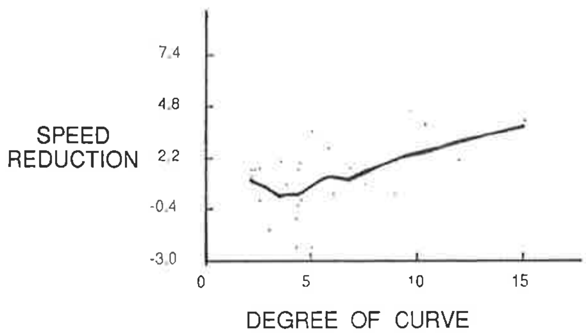


FIGURE 5 Speed reduction versus degree of curve: favorable curve group.

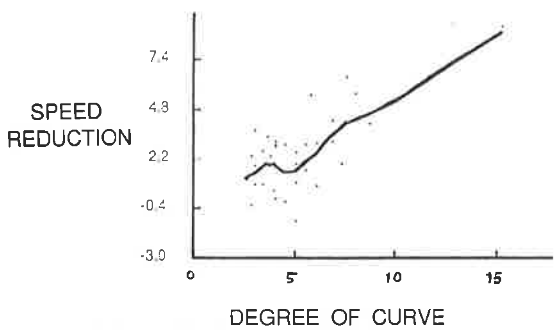


FIGURE 6 Speed reduction versus degree of curve: unfavorable curve group.

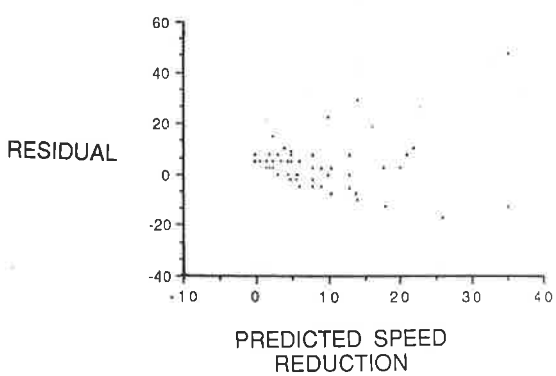


FIGURE 7 Predicted speed reduction versus residual.

Locally Weighted Regression

Because of the extreme amount of dispersion in the data, LOWESS was employed. This method of regression is a nonparametric approach that does not assume constant error variance of the dependent variable as does the method of least squares regression.

The LOWESS technique fits a line within certain groups of points and then joins the lines for each group to form a curve. This technique gives only the shape of the curve. It does not give parameter estimates. Illustrative LOWESS plots of speed reduction versus degree of curve for the favorable and unfavorable groups of curves can be seen in Figures 5 and 6, respectively. Plots for the ungrouped curves (i.e., combining favorable and unfavorable geometric combinations) are depicted in Figures 2–4 for speed reduction.

In view of the various LOWESS plots, an unclear relationship exists between the values of the five operational variables and degree of curve values below 5 degrees. However, a number of the graphs depict a linear relationship when considering curves of 5 degrees and higher, particularly speed reduction and edgeline encroachment rates for inside lanes.

There are relatively few sharp curves (at least 10 degrees) in the sample. The corresponding LOWESS plots in this region would not be as reliable as those for milder curves.

SUMMARY AND CONCLUSIONS

In conclusion, several relationships bear mentioning. Average speed reduction and edgeline encroachments on curves to the right appear to be positively associated with degree of curve for curves about 5 degrees. As curves become sharper, there is a proportionally greater increase in speed reduction and edgeline encroachments for curves to the right. Centerline encroachments on curves to the left also increase more drastically than those on curves to the right.

These results on operational measures support results of the accident analyses presented by Zegeer et al. (*1*). For example, degree of curve is clearly the geometric feature that most adversely affects both accidents and vehicle operations on horizontal curves, with sharper curves producing significantly increased rates of accidents, as well as high rates of speed reductions and vehicle encroachments. The greater incidence of edgeline encroachments for curves to the right combined with increased centerline encroachments for curves to the left supports the contention of drivers undercutting the curve (*2*). This practice can result in run-off-road crashes on curves to the right or head-on and opposite-direction–sideswipe accidents with oncoming motorists on curves to the left.

The results of the operational analysis support the predominance of single-vehicle crashes (i.e., fixed-object and rollover) and opposing multivehicle crashes (i.e., head-on and opposite-direction–sideswipe), which have been found in various curve studies (*1*,*2*) to be overrepresented on curves when compared with tangents. The sharpness of curves in particular is associated with greater speed reductions and encroachments over the centerline and edgeline. To a lesser degree, wide shoulders and less severe grades were associated with fewer operational problems. Such results reinforce the importance of various design features in curve safety and operations.

RECOMMENDATIONS FOR CURVE DESIGN AND UPGRADING

The study findings are of interest in determining appropriate, cost-effective countermeasures to existing curve problems. They are also useful in evaluating the adequacy of current curve design policy and in formulating guidelines for design of high-speed curves.

Of course, countermeasure effectiveness and design standards for curves are largely oriented toward safety. Operational measures of effectiveness are valuable to the extent that they support or clarify what is known about the relationships between curve design features and accidents.

This section presents findings on the subject of curve design for existing as well as new highways. These findings are based not only on the operational studies reported here, but also on separate analyses of accident and roadway data bases (*1*). These analyses, in turn, build on previous research (*2*,*4*).

The operational analyses clearly represent only a few pieces of the puzzle describing the safety and operational characteristics of highway curves. Nonetheless, the findings are important in that they tend to support other key research on safety.

In general terms, it can be concluded that highway curves present special problems to drivers. Curves require drivers to adjust their speed and path. When curves are particularly sharp and when no transition curvature is provided, driver and vehicle behavior problems ensue. Furthermore, when such curves occur in concert with other geometric problems (e.g., steep grades, narrow lanes, narrow shoulders, and hazardous roadsides), accidents are inevitable.

Previous work has led to the following conclusions about treatment of existing curve problems:

1. Flattening an existing sharp curve offers the greatest potential (relative to other improvements) for reducing crash frequency.
2. Roadway widening on curves is effective in producing lower accident rates.
3. The presence of spiral transitions and adequate superelevation is associated with small, but significant, reductions in accidents.
4. Various roadside improvements (e.g., clearing trees near the road, relocating utility poles, and flattening sideslopes) are particularly beneficial at curves.

Cost-Effective Treatments of Existing Curves

Not every sharp curve represents a safety or operational problem, and not every high-accident curve can be treated by geometric or traffic control countermeasures. Although research findings provide insights into curve countermeasures, site-specific study is clearly mandated. Each location is unique in its constraints, physical conditions, and operational characteristics. The study of existing accident patterns, an evaluation of site geometric and roadside conditions, and observations of driver behavior are necessary to adequately identify appropriate treatments at a particular curve site.

Generally, countermeasures fall into three categories: (a) complete reconstruction; (b) physical rehabilitation or partial reconstruction; and (c) low-cost spot improvements, such as signing, marking, and delineation.

Curve reconstruction represents the most costly, but potentially the most effective, means of reducing severe curve accidents. Curve reconstruction may involve flattening of the curve, widening of lanes and shoulders, new pavement, improved roadside, and the addition of a spiral transition curve. The feasibility or cost-effectiveness of total curve flattening and reconstruction depends largely on site-specific conditions. The availability and cost of right-of-way, vertical alignment requirements, environmental impacts, and local access changes would influence any decision to reconstruct a curve.

Curve rehabilitation and partial reconstruction are typically less costly measures than curve flattening or roadway widening and may be highly effective in treating existing curves. Foremost among these is removal of roadside hazards within the area of influence of the curve. Tree removal, utility pole relocation, sideslope flattening, and other improvements may be cost-effective at relatively low traffic volume levels. Resurfacing of the curve to improve skid resistance is also a potential solution. Resurfacing projects can be used to improve the superelevation in the curve, adjust the superelevation transition, pave the shoulder through the curve, clear roadside obstacles, and eliminate pavement edge drop-off conditions. All of these measures can be implemented within existing rights-of-way and with relative ease. The effectiveness of a package of curve rehabilitation countermeasures would, of course, depend on the particular site. TRB *Special Report 214* (*8*) provides useful information about resurfacing, restoration, and rehabilitation.

Such improvements as signing, marking, and delineation are intuitively appealing because of their low cost and ease of implementation. Advance warning signs, centerline and edgeline markings, and special delineation schemes have been tested at high-accident locations. Special attention to signing and marking is important along any highway, particularly at sharp curves. It is clear, however, that the addition of signing, marking, and delineation cannot be expected to solve a safety problem on a poorly designed curve. At the same time, proper signing, marking, and delineation in accordance with the *Manual on Uniform Traffic Control Devices* (*9*) are essential when treating hazardous curves in conjunction with other improvements (e.g., clearing roadsides, widening the roadway, paving the shoulder, flattening the curve, or improving the superelevation). Even if construction or reconstruction of a poorly designed curve is not possible, substandard signing, marking, and delineation should still be improved on hazardous curves.

Design Guidelines for New Highway Sections

Most highway design in the United States is governed by AASHTO procedures, criteria, and design values (*10*). Research from all recent studies on horizontal curves suggests that application of some specific design guidelines would significantly improve the overall quality of horizontal curve design.

For example, designers should provide for consistent roadway sections. By avoiding sharp, isolated curves and by maintaining consistency in the design of superelevation, roadway width, and other features on curves, designers can minimize the element of surprise and better accommodate the difficult speed and path transition behavior required of drivers.

Designers should avoid large central angles wherever possible. Such angles force designers to choose between long curves or sharp curves, both of which present problems. On long curves driver exposure to curve tracking problems (centerline and edgeline encroachments) is longer. On sharp curves tracking and speed transition problems are more severe. A suggested criterion is to avoid central angles greater than 45 degrees.

Designers should minimize the use of controlling curvature (i.e., maximum allowable curvature for a given design speed). Many designers tend to view all curves as equally safe within a given design speed. However, milder curves operate better and tend to have better accident histories. Where controlling curvature is used, designers should pay extra attention to the roadside design (particularly on curves to the left).

Designers should routinely use spiral transition curves, particularly for controlling curves on highways with high design speeds (e.g., 60 mph or greater). The safety effects of spirals have been demonstrated (*1*) and operational benefits of spirals have also been found (*2*). Again, adequate transition design is particularly critical on higher speed alignment.

Designers should routinely provide high-quality roadside designs, particularly on sharper curves. Wider shoulders, flatter sideslopes, and greater roadside clear zones in these areas are essential design features.

Designers should avoid locating other potentially hazardous features at or near horizontal curves, in recognition of driver difficulty in tracking curvature. Features to avoid include intersections, narrow bridges, major cross-section transitions, and driveways. Another potentially hazardous feature is severe reverse curvature with curves in opposing directions separated only by a short tangent alignment.

Designers should provide adequate pavement and shoulder condition, particularly on sharper curves where lateral acceleration and function demand are the greatest. Increasing pavement skid resistance is often an essential curve improvement, especially for curves that have skidding accidents during wet pavement conditions. On highways with unpaved shoulders, consideration should be given to paving the shoulders at the sharper curves. Vertical curvature should be provided such that more than the minimum stopping sight distance is available throughout the curve.

Finally, designers should use an adequate amount of superelevation on all curves.

ACKNOWLEDGMENTS

This paper was based on an FHWA study entitled, *Cost-Effective Geometric Improvements for Safety Upgrading of Horizontal Curves*. Appreciation is expressed to George Pilkington and Justin True for their input.

REFERENCES

1. C. Zegeer, R. Stewart, D. Reinfurt, F. Council, T. Neuman, E. Hamilton, T. Miller, and W. Hunter. *Cost-Effective Geometric Improvements for Safety Upgrading of Horizontal Curves*. Report FHWA-RD-90-021. FHWA, U.S. Department of Transportation, 1990.
2. J. C. Glennon, T. R. Neuman, and J. E. Leisch. *Safety and Operational Considerations for Design of Rural Highway Curves*. Report FHWA-RD-86-035. FHWA, U.S. Department of Transportation, 1985.
3. B. E. Jennings and M. J. Demetsky. Evaluation of Curve Delineation Signs. In *Transportation Research Record 1010*, TRB, National Research Council, Washington, D.C. 1985.
4. K. W. Terhune and M. R. Parker. *An Evaluation of Accident Surrogates for Safety Analysis of Rural Highways*. Report FHWA-RD-86-127. FHWA, U.S. Department of Transportation, 1986.
5. T. K. Datta et al. *Accident Surrogates for Use in Analyzing Highway Safety Hazards*. Report FHWA-RD-82-103-105. FHWA, U.S. Department of Transportation, 1983.
6. T. H. Rockwell, J. Malecki, and D. Shinar. *Improving Driver Performance on Rural Curves Through Perceptual Changes*. Ohio State University, Columbus, 1975.
7. T. H. Rockwell and J. C. Hungerford. *Use of Delineation Systems to Modify Driver Performance on Rural Curves*. Ohio State University, Columbus, 1979.
8. *Special Report 214: Designing Safer Roads: Practices for Resurfacing, Restoration, and Rehabilitation*. TRB, National Research Council, Washington, D.C., 1986.
9. *Manual on Uniform Traffic Control Devices for Streets and Highways*. FHWA, U.S. Department of Transportation, 1988.
10. Task Force on Geometric Design. *A Policy on Geometric Design of Highways and Streets—1984*. AASHTO, Washington, D.C., 1984.

Publication of this paper sponsored by Committee on Vehicle User Characteristics.

Effect of Pavement and Shoulder Condition on Highway Accidents

JOSEPH CRAUS, MOSHE LIVNEH, AND ILAN ISHAI

The effect of pavement surface condition on traffic accidents was investigated as part of an Israeli Public Works Department survey of the feasibility of investing in the maintenance of the interurban road network in Israel. An analysis of this problem for the whole network attempted to correlate the state of the pavement surface with conditions of safety. The results of the investigation did not indicate any unilateral correlation between these two parameters. In contrast, similar tests that concentrated on specific road sections indicated that an increase in pavement surface grading might improve or worsen traffic safety, depending on the geometric and traffic characteristics of the section under investigation. Additional tests conducted in the context of this survey concerned the state of the pavement shoulders and the slipperiness of the pavement. According to the Israeli data, it is possible to decrease the total number of accidents (nonintersection accidents) on the interurban road network by approximately 7.5 percent if the skid coefficient of the pavement surface, as measured by the Mu-Meter, is greater than 37. Similarly, it is possible to diminish the number of such accidents by about 8 percent if the shoulder's surface is in good condition and the width of the shoulders is no less than 2 m. The findings of this investigation indicate that black spots or black section analyses are required rather than an analysis of the whole network to determine the pavement and shoulder condition on highway accidents. An application of the black spots or black sections concept facilitates the achievement of an optimal level of economic feasibility. Thus, for example, antiskid treatments provided at those sections where there is a high rate of accidents of this type, or treatment of the pavement shoulders and their widening at those sections where there is a high rate of accidents caused by a state of disrepair, can lead to high benefit-cost values.

One of the issues investigated as part of an Israeli Public Works Department survey (*1*) of the feasibility of investing in the maintenance of interurban roads was the influence of the state of the pavement surface and shoulders on the rate of accidents. Obviously, any decrease in the rate of accidents constitutes part of the benefits obtained by investing in routine or major maintenance of the road network.

Naturally, the existence of such a correlation between the state of the pavement surface and traffic safety seems reasonable. Improving the pavement surface condition by overlaying its surface is conducive to a decrease in the number of accidents. As reported by TRB (*2*), an overlay on two-lane roads leads to a decrease of 12 percent in the number of accidents, most of them being wet accidents. Additionally, some American sources, cited by Hakkert (*3*), report that resurfacing reduces wet accidents by 12 to 33 percent. According to one source (*4*), the cost of accidents increases by 10 percent for pavements in a deteriorated condition. Haas and Hudson (*5*) state that the cost of accidents varies according to the type of road and the state of the pavement. For example, for two-lane roads the cost difference in accidents between a good pavement with a present serviceability index (PSI) equal to 4 and a bad pavement with a PSI equal to 2.5 is 13 percent. In the last two references (*4,5*), no emphasis is given to accidents on wet pavements. In contrast to these sources, others point out that a decrease in the number of accidents on dry pavement as a result of an improvement in the pavement state is balanced by accidents caused as a consequence of increased speeds. For example, TRB (*6*) states that overlaying activity increases the initial accident rate on a dry road by 10 percent and decreases the initial rate of accidents on wet roads by 15 percent. Thus, summation of these two rates eventually cancels out the effect of overlaying on the change in the total rate of accidents, when the total initial increase is about 5 percent.

As a result of these findings, an investigation was conducted to attempt to determine whether there is a decrease in the number of dry-road accidents in the interurban road network in Israel as a result of improved pavement surfaces and to present specific examples of the possible rate of change in such accidents for some specific cases. Furthermore, the investigation was required to estimate the extent of the decrease in wet-road accidents after proper antiskid overlaying and to estimate the rate of the decrease in road-shoulder accidents after improvement.

Itemization of the issues connected with these tasks is presented in the following sections.

GENERAL CASE

To examine the influence of pavement state on the incidence of accidents, the Israeli roads network was classified into five pavement state groups according to the Washington State Department of Transportation criteria (*7*), as presented in Table 1. A rating of 100 in this table signifies a pavement in excellent state (pavement immediately after paving or over-

TABLE 1 CLASSIFICATION OF ROADS NETWORK ACCORDING TO PAVEMENT SURFACE CONDITION

Group No.	Range of Pavement Rating (PR) Values
1	90-100
2	70-90
3	50-70
4	30-50
5	0-30

Transportation Research Institute, Technion-Israel Institute of Technology, Technion City, Haifa 32 000, Israel.

laying), whereas a rating of 0 signifies the worst state of pavement surface (a totally deteriorated pavement).

The average accident rate per vehicle-kilometer was computed for each group (1983 data). These data are shown in Figure 1. No visible influence of the pavement state on the rate of accidents is indicated.

Additionally, the Kruskal-Wallis H test has been used to test whether the accident rates of the five groups are from the same populations. In this test the H value is calculated according to the following equation:

$$H = \frac{12}{N(N+1)} \left(\sum_{i=1}^{5} \frac{R_i^2}{N_i} \right) - 3(N+1) \quad (1)$$

where

N = number in all groups combined,
N_i = number in i groups, and
R_i = sum of ranks in the ith group.

To rank the accident rates shown in Figure 1, these rates should be normalized for an average traffic volume (AADT). This procedure is done using the following regression lines (shown in Figure 1):

Single Carriageway

$A = 0.454 - (0.016/1{,}000) \times \text{AADT}$

$R^2 = 0.233$

$N = 20$ (2)

Dual Carriageway

$A = 0.128 - (0.001/1{,}000) \times \text{AADT}$

$R^2 = 0.154$

$N = 10$ (3)

Single and Dual Carriageway

$A = 0.404 - (0.010/1{,}000) \times \text{AADT}$

$R^2 = 0.438$

$N = 30$ (4)

The normalized values obtained by the three regression lines (mean and standard deviation only) are presented in Table 2.

For the detailed normalized values of the single-carriage way accident rates, $H = 5.257$. This value of H is interpreted as chi-square with the number of samples minus 1 degree of freedom. For the given problem $d_f = 4$ and the tabulated value of chi-square, is 9.488 when $\alpha = 0.05$. Because $9.488 > 5.257$, H is not significant; that is, all the values of accident rates for the five groups of single carriageway are from the same population. For the dual-carriageway roads, a similar calculation indicates that the calculated value of H is 3.055, which is not significant ($9.488 > 3.055$). The same conclusion applies to both single- and dual-carriageway roads for which the calculated value of H is 2.70 ($9.488 > 2.70$).

The conclusion that pavement state does not influence the rate of accidents is only applicable in the general case of the whole network. However, it cannot be established that the improvement in pavement state does not decrease (or even increase) the rate of accidents in specific cases. This issue is discussed in the following section. The data in Figure 1 also indicate that, for a certain intermediate range of AADT (between 10,000 and 20,000 AADT), single-carriageway roads do not necessarily have higher accident rates than dual-carriageway roads, as indicated by the scatter of points in that region.

Similar conclusions were reached in Israel by other investigators (*8*). Figure 2 shows accident rates on single carriageways with high volumes (more than 10,000 AADT). Although

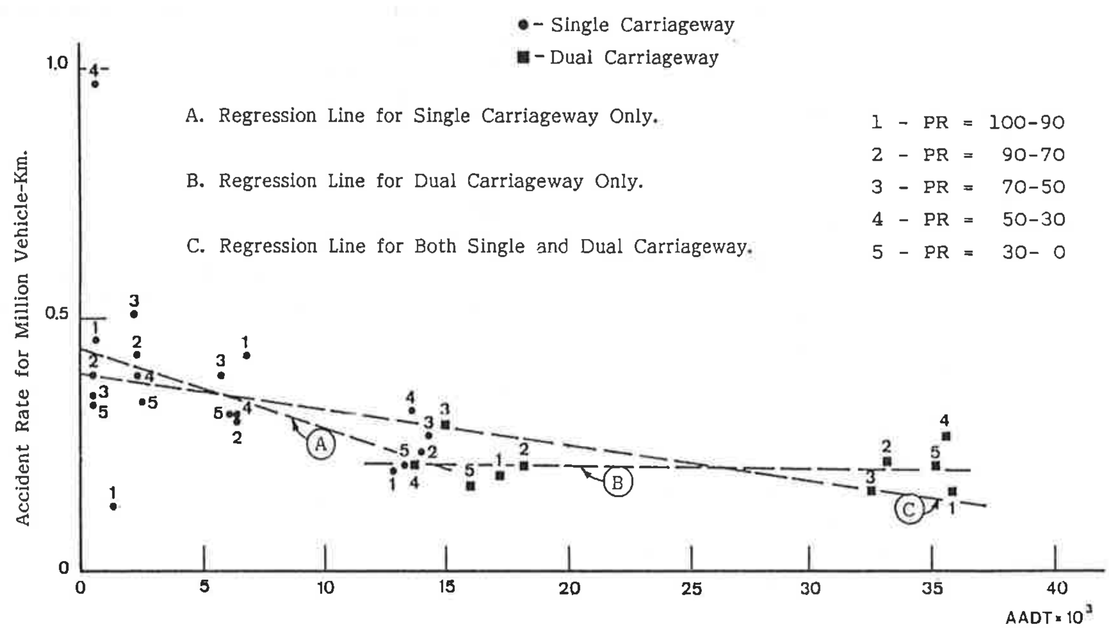


FIGURE 1 Relationship between yearly accident rates per million vehicle-kilometers and AADT values as a function of the pavement surface state (*1*).

TABLE 2 NORMALIZED ACCIDENT RATES PER 10^6 vehicle-km FOR THE FIVE GROUPS OF PAVEMENTS

Range of PR	Single Carriageway corrected for mean AADT of 5685		Dual Carriageway corrected for mean AADT of 24660		Single & Dual Way corrected for mean AADT of 12010	
	Mean	st. dev.	Mean	st. dev.	Mean	st. dev.
90-100	0.302	0.160	0.090	0.001	0.237	0.134
70- 90	0.304	0.099	0.108	0.011	0.266	0.060
50- 70	0.382	0.078	0.112	0.035	0.289	0.079
30- 50	0.498	0.260	0.118	0.032	0.371	0.246
0- 30	0.295	0.037	0.115	0.004	0.230	0.079

rates are scattered over a wide range, an average value of 0.5 accident per million vehicle-kilometers can be observed. Figure 3 shows values of accident rates on dual-carriageway roads. Rates vary from a high value of 1.0 accident per million vehicle-kilometers to a low value of 0.1 accident per million vehicle-kilometers, according to the specific conditions of the roads. This finding confirms the previous conclusion that dual-carriageway roads do not necessarily have lower accident rates.

SPECIFIC CASES

In contrast to the general investigation presented in the preceding section, this section illustrates a specific investigation for two practical examples. The first site examined is Road 4, located between the 131st and 148th kilometers (Ra'anana intersection to the Hasharon intersection). This site is a four-lane road with a dual carriageway. A 40 percent decrease in the total rate of nonintersection accidents has been reported as a consequence of pavement surface overlaying (see Table 3). Most of the accidents in this section are dry-road accidents. There was a similar rate of decrease (38 percent) in intersection accidents. However, this decrease must be ascribed to the changes in traffic arrangements that resulted in some of the intersections, especially at the Hasharon intersection, rather than to the overlaying work itself.

Similarly, another road section can be examined in which overlaying might increase the rate of accidents as a consequence of improving the riding conditions of a road with low geometric standards. The road examined is the 13.7-km-long Road 806 (Ma'ar-Elaboon). This road is situated in a mountainous region with low geometric standards, and its surface is extremely deteriorated. The rate of accidents on this road from 1982 through 1987 is an average of approximately 0.3 accident per million vehicle-kilometers of travel, with an average traffic volume of 2,000 vehicles per day. This accident rate is low and corresponds to that of a road with high standards. This rate exists in the road section under consideration solely because of bad travel conditions. Any improvement of these conditions through overlaying that is not accompanied by geometric improvements will increase the rate of dry-road accidents. Various authors have presented estimations of the increase in the number of accidents as a consequence of pavement resurfacing without improving the geometry of the road. On the one hand, the improvement in pavement surface invites speeding; on the other hand, a low geometric standard contributes to an increase in the number of accidents under

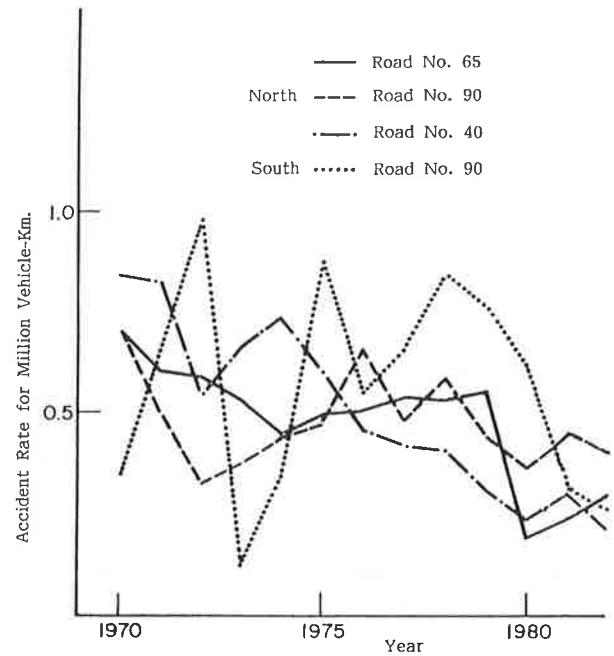


FIGURE 2 Accident rates per million vehicle-kilometers on major interurban roads in Israel (1970–1982): single carriageway (*8*).

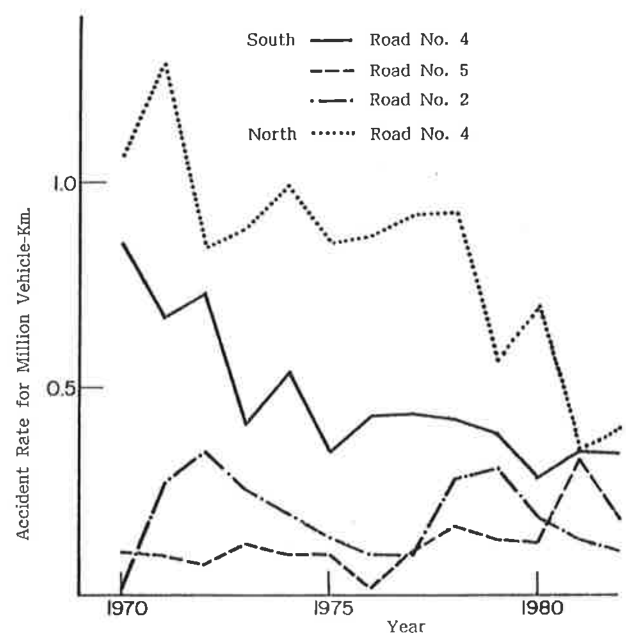


FIGURE 3 Accident rates per million vehicle-kilometers on major interurban roads in Israel (1970–1982): dual carriageway (*8*).

TABLE 3 COMPARISON OF ACCIDENTS BEFORE AND AFTER OVERLAYING ON ROAD 4 BETWEEN THE RA'ANANA AND HASHARON INTERSECTIONS

Place	After overlaying	Before overlaying	
	Year 1986	Year 1983	Year 1982
Raanana intersection North	5	11	5
Basra intersection	1	--	1
Dror intersection	8	5	2
Even Yehuda intersection	1	--	1
Pardesia intersection	1	1	1
Hasharon intersection	2	12	13
Total intersection	18	29	23
Total elsewhere	15	25	23
Total accidents	33	54	46

conditions of increased speed. On the basis of before and after studies conducted by the Transport and Road Research Laboratory (TRRL) (*9*) on English roads, it is estimated that this increase in the accident rate could amount to 65 percent. It has also been stated that "pavement resurfacing without other highway improvements will have a small negative effect on safety on most roads" (*6*).

These two examples present opposing phenomena, yet each example has its own unique explanation. On the dual-carriageway road, most of the traffic concentrated in the left lane before overlaying because of deterring aberrations in the surface of the right lane. Obviously, such a state could encourage dangerous driving maneuvers, which in turn could engender the proliferation of accidents. After overlaying of the pavement, the traffic resumed its use of both lanes of travel with a natural distribution, and the redundant and dangerous driving maneuvers disappeared.

For the road with low geometric standards, the pavement surface that is subject to improvement causes drivers to increase their speeds and drive less cautiously after overlaying than they would on a road with low standards. This behavior could increase the number of dangerous maneuvers and thus, of course, the rate of accidents. Before overlaying, the deteriorated pavement surface increases the caution of the drivers and thus led to a lower rate of accidents.

These two examples illustrate that the effect of the pavement state on traffic safety is correlated with the unique characteristics of the case under investigation and cannot be established for the whole network.

WET ACCIDENTS

In rainy weather, road accidents can be engendered by the following causes:

- Accidents stemming from the same causes that generate accidents on dry pavements,
- Accidents stemming from low visibility in rainy weather, and
- Accidents stemming from skidding on a wet pavement.

Obviously, antiskid overlaying of the pavement surface can only decrease the rate of skid accidents. To calculate the extent of this accident rate, the following expression should be applied:

$$A_{sw} = m\left(A_w - \frac{A_w}{n+1}\right) \times 100 \quad (5)$$

where

A_{sw} = proportion of skid accidents out of the total of all accidents on the road,
A_w = proportion of wet-pavement accidents out of the total of all accidents on both wet and dry pavements,
n = relative frequency of accidents on a wet pavement compared with accidents on a dry pavement, and
m = proportion of skid accidents out of the total of wet-pavement accidents.

The value of n can be calculated from the following facts. First, according to Nedavia (*10*), the time during which the pavement is wet is estimated to be 3 percent of the total year. This estimation is based on data obtained from the meteorological service to which were added the number of hours that the roads remain wet after the rain. Because of the lack of precise data on actual traffic volumes during wet periods, it was assumed that the traffic volume is proportional to the condition of the pavement, that is, that 97 percent of the traffic takes place on dry pavement surfaces and that 3 percent takes place on wet pavement surfaces. Second, Table 4 indicates that the percentage of accidents on a wet pavement out of the total number of accidents is 12 percent. Thus, the relative frequency (n) of accidents on wet pavement compared with accidents on dry pavement that are caused by traffic throughout the network is

$$n = \frac{n_w}{n_d} = \frac{12/0.03}{88/0.97} = 4.4 \quad (6)$$

where

n_w = frequency of accidents on wet pavement (in the preceding case the rate of accidents is 12 percent and the traffic volume is 3 percent of the total traffic passing through the network in vehicle-kilometers); and

TABLE 4 STATE OF ACCIDENTS ON ISRAELI ROAD SECTIONS DURING 1983–1986

Type of Road	1986		1985		1984		1983	
	Total Accd.	Wet Accd.	Total Accd.	Wet Accd.	Total Accd.	Wet Accd.	Total Accd.	Wet Accd.
Dual carriageway	672	81	554	55	546	66	645	82
Main road	928	126	811	87	842	98	830	102
Regional road	332	43	311	38	282	28	311	39
Local road	84	10	57	0	59	3	63	6
Undefined road	18	2	9	0	11	0	21	0
Total	2034	262	1742	180	1740	195	1870	229

n_d = frequency of accidents on dry pavement (the frequency of accidents is 88 percent, and the traffic volume is 97 percent of the total traffic passing through the network in vehicle-kilometers).

The result obtained is that the frequency of accidents on wet pavement is 4.4 times that of accidents on dry pavement (1.26 wet accidents per million vehicle-kilometers in 1983 compared with 0.28 dry accident per million vehicle-kilometers in that year). This finding indicates the high frequency of accidents on wet pavement. In other countries, such as France, this frequency is less than 2, which indicates that the pavement surface resistance to skid is at a lower standard in Israel.

This finding leads to the estimation that the percentage of skid accidents in Israel out of the total accidents on wet pavement (m) is 80 percent, in contrast to lower values in other countries (such as 55 percent for England). Hence, the percentage of skid accidents out of the total number of road accidents that can be prevented through proper overlaying is

$$A_{sw} = 0.8\left(0.12 - \frac{0.12}{5.4}\right) \times 100 = 7.5 \text{ percent} \qquad (7)$$

Another method of estimating the rate is based on the data in Figure 4. This figure shows the percentage of wet-pavement accidents compared with the total number of accidents, depending on the relationship between the length of those sections in which the friction value SN on the Mu-Meter is less than 37 (sections with low skid resistance) and the total length of the sections. According to this figure, the percentage of accidents is 9 percent when all the sections examined have an appropriate skid coefficient and 16 percent when they do not. Hence, increasing the skid coefficient decreases the rate of wet accidents from 16 to 9 percent. Thus,

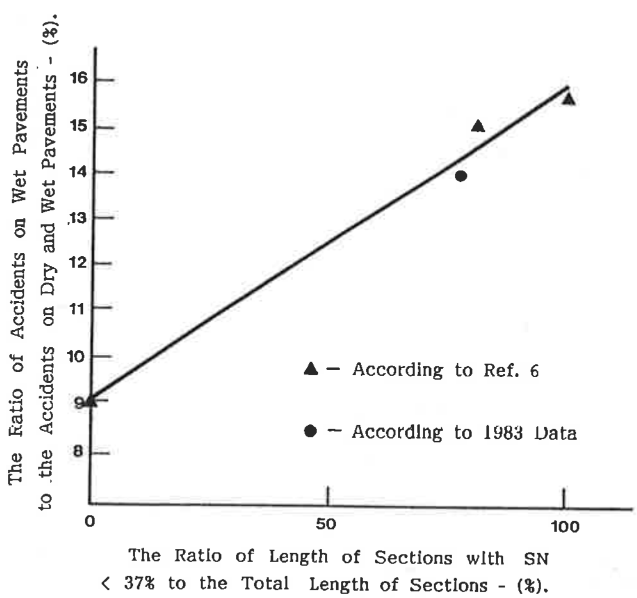


FIGURE 4 Relationships between the relative length of a section with low SN value and wet accident rates (*1*).

$$\frac{W}{W + D} = 0.16 \qquad (8)$$

$$\frac{W - dW}{W - dW + D} = 0.09 \qquad (9)$$

where

W = number of accidents on wet pavement,
D = number of accidents on dry pavement, and
dW = decrease in number of accidents on wet pavement as a result of raising the skid coefficient from an SN value lower than 37 to one higher than 37.

Equations 8 and 9 give D/W equals 5.25, dW/W equals 0.48, and dW/(W + D), that is, A_{sw} equals 7.7 percent. This result is practically identical to that presented at the beginning of this section (7.5 percent) and constitutes proof that the percentage of skid accidents out of the total number of accidents on wet pavement is 80 percent.

SOUNDNESS OF ROAD SHOULDERS

In considering the effect of the shoulders on traffic safety, it is appropriate to quote the following excerpt (*6*):

> Wide lanes and shoulders provide motorists with increased opportunity for safe recovery when their vehicles run off the road (an important factor in single vehicle accidents) and increased lateral separation between overtaking and meeting vehicles (an important factor in sideswipe and head-on accidents). Additional safety benefits include reduced interruption from both emergency stopping and road maintenance activities, less wear at the lane edge, improved sight distance at critical horizontal curves, and improved roadway surface drainage.

To learn the influence of shoulder state and width on the rate of accidents in Israel, 1983 data pertaining to nonintersection accidents on interurban roads in Israel are presented in Table 5. Both the total number of accidents and only single-vehicle accidents are examined according to sound and unsound shoulders (as determined by the police examiner's classification), where the width of the shoulders is < 2 m and where it is ≥ 2 m.

On the basis of the many local observations and studies, it has been found that the rate of single-vehicle accidents increases in segments with unsound shoulders (narrow shoulders, shoulders in a deteriorated state, soft shoulders, etc.). When a single vehicle veers off the road onto an unsound shoulder, it is difficult for the driver to safely return to the road. There is often a loss of steering control, which can lead to a single-vehicle accident. This trend can also be observed in Table 5.

According to the quotation, the ratio between the number of single-vehicle accidents and the total number of accidents can be viewed as an indication of the extent of the shoulders' efficiency. Obviously, as this ratio increases, it can be assumed that the efficiency of the shoulders decreases, when the efficiency itself is correlated with the soundness and width of the shoulders. Indeed, for the entire network the relationship for unsound shoulders is 66 percent. This relationship is smaller

TABLE 5 NONINTERSECTION ACCIDENTS ON INTERURBAN ROADS IN 1983

State and Width Of Shoulders	Length in km	Number of Accidents		% Single Vehicle Accidents
		Total	Single Vehicle	
General Total				
Bad	325	50	33	66.0
Medium and Good	3555	1694	584	34.5
Total	3880	1744	617	35.4
Less Than 2 m.				
Bad	217	26	20	76.9
Medium and Good	1182	247	91	36.8
Total	1399	273	111	40.7
2 m. and more				
Bad	108	24	13	54.2
Medium and Good	2370	1447	493	34.1
Total	2478	1471	506	34.4

by almost half (34 percent) for sound shoulders. Figure 5 and Table 5 both indicate that the width of the shoulders is also an important factor in the issue of travel safety. The ratio between the number of single-vehicle accidents and the total number of accidents increases to 77 percent for unsound shoulders less than 2 m wide and to 54 percent for unsound shoulders 2 m wide or more.

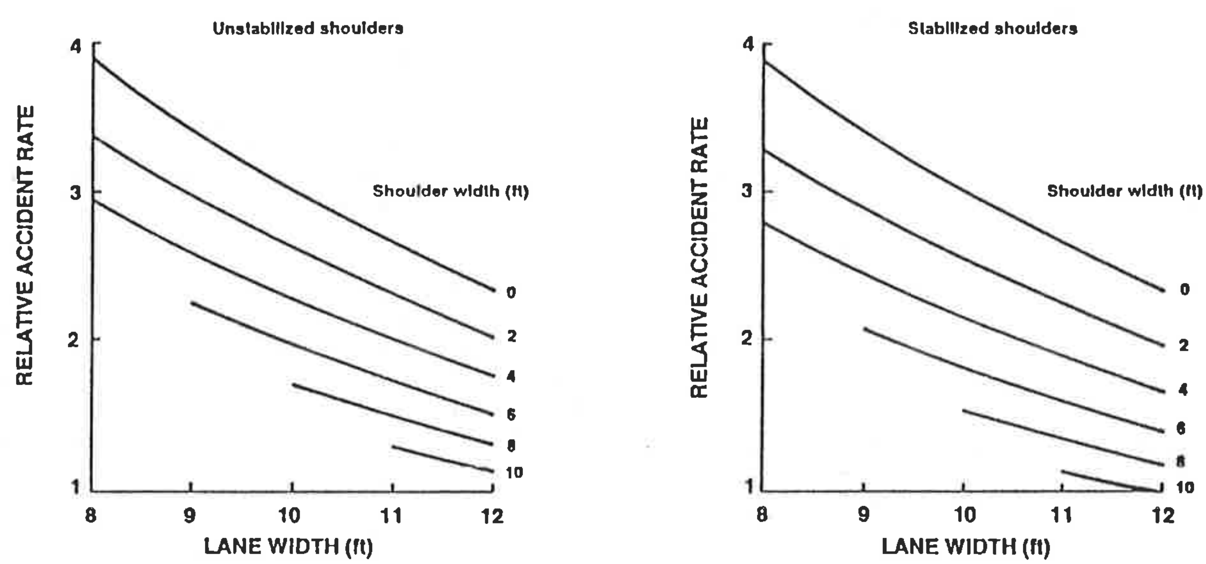
According to the data in Table 5, it is possible to compute the percentage of accidents out of the total number of accidents (A_s) that can be prevented if all the shoulders are sound and of a width of 2 m. This calculation is as follows:

$$\frac{0.354 - A_s}{1 - A_s} = 0.341 \tag{10}$$

This calculation leads to an A_s value equal to 2 percent. It should be assumed that the actual value would have been higher if the classification of the shoulder had been conducted to study their effect on riding safety rather than to fill out an accident report. Thus, for example, if the real ratio between single-vehicle accidents and the total number of accidents in a road section of sound shoulders with a width of ≥ 2 m is 30 percent (rather than 34.1 percent), the value of A_s increases to 8 percent. This latter value seems more realistic in light of the data shown in Figure 5.

SUMMARY AND CONCLUSIONS

The data presented lead to the conclusion that the issue of road safety should be viewed not in general terms (i.e., systems concepts) but in the specific terms of black spots or black sections. Upgrading the state of the pavement surface does not always decrease the rate of accidents; therefore, each case must be examined separately. General terms that would aspire to be identically pertinent to every section of the network cannot be used. Even the question of widening a single-carriageway road into a dual-carriageway road cannot be assessed using uniform, generalized concepts. There are dual-carriageway roads where the accident rate is far lower than that on single-carriageway roads, but there are also dual-carriageway roads with a far higher accident rate. Hence, it is concluded that additional factors, such as geometric and environmental elements dictate the possible change that can be effected in the accident rate through widening the road to a dual-carriageway one. This question, too, must be specifically examined in each particular case (*8*).



NOTES: Accident relationship covers single-vehicle, sideswipe, and opposite-direction accidents on two-lane rural highways. Relative accident rate is defined as a multiple of the accidents per million vehicle miles for 12-ft lanes and 10-ft stabilized shoulders.

FIGURE 5 Normalized relationship between accidents and lane and shoulder conditions (*6*).

Increasing the skid coefficient of the pavement surface reduces the wet-accident rate. Here, a general decrease rate can be established for the whole network (7.5 percent of all accidents in road sections). However, from the point of view of economic feasibility, discussion should obviously refer only to black spots or black sections. There is no doubt that high economic feasibility can be achieved only when those sections in which there is a high incidence of skid accidents are treated. Hence, in Israel the following rule serves as the minimal criterion required for justifying the application of an antiskid asphalt layer: An application of asphalt layer is recommended if there have been at least 6 accidents during the preceding 3 years on a road where the percentage of wet-accidents out of the total number of accidents is greater than 30 percent or if there have been at least 12 accidents during the preceding 3 years on a road where the percentage is 25 percent.

Increasing the soundness of the shoulders decreases the incidence of accidents in which the shoulders are involved. Here, too, treatment is only worthwhile on those sections with a high incidence of accidents.

In summary, the influence of the state of the pavement surface and shoulders on riding safety has been addressed. According to the Israeli data, the total number of accidents (nonintersection accidents) on the interurban roads network can be decreased by approximately 7.5 percent if the skid coefficient of the pavement surface, as measured by the Mu-Meter, is ≥ 37. Similarly, it is possible to diminish the number of such accidents by approximately 8 percent if the perfect soundness of the shoulders is ensured and their width is increased to 2 m or more.

The value of 8 percent has been chosen as more realistic than the calculated value of 2 percent. The latter value is extremely conservative, stemming from the routine police examiner's classification method for the state of the shoulders. A combination of the two activities might annually prevent 250 to 300 accidents in Israel.

ACKNOWLEDGMENT

This paper is based partly on data obtained from work carried out for the Israeli Public Works Department.

REFERENCES

1. L.C.I., Cambridge Systematics, Inc., and B. Ellenbogen. *Pavement Management System—Stage A: Economic Engineering Survey of the Efficacy of Maintenance Investment of Interurban Road Network in Israel* (in Hebrew). Report submitted to the Israeli Public Works Department, Jan. 1988.
2. *NCHRP Report 162: Methods for Evaluation of Highway Safety Improvements*. TRB, National Research Council, Washington, D.C., 1975.
3. A. S. Hakkert. *Determining the Priority of Road Safety Improvements*. Technical Report RY/6183. National Institute for Transport and Road Research, Council for Industrial and Scientific Research, Pretoria, South Africa, 1983.
4. *Uses Costs*. TVH Finland, 1985.
5. R. Haas and W. R. Hudson. *Pavement Management Systems*. McGraw-Hill, New York, 1978.
6. *Special Report 214: Designing Safer Roads. Practices for Resurfacing, Restoration and Rehabilitation*. TRB, National Research Council, Washington, D.C., 1987.
7. T. L. Nelson and R. V. Leclerc. *Development and Implementation of Washington State's Pavement Management System*. Report WA-RD-50.1. Washington State Department of Transportation, Olympia, 1983.
8. A. S. Hakkert. *Analysis of Road Accident Trends in Israel* (in Hebrew). Report 87-118. Transportation Research Institute, Technion-Israel Institute of Technology, Haifa, 1987.
9. Transport and Road Research Laboratory. *Research on Road Traffic*. Her Majesty's Stationery Office, London, England, 1965.
10. D. Nedavia. *Research on the Slipperiness of Wet Asphalt Pavements in Israel* (in Hebrew). Report 48. Materials and Research Department, Israeli Public Works Department, 1981.

Publication of this paper sponsored by Committee on Traffic Records and Accident Analysis.

Head-On, Left-Turn Accidents at Intersections with Newly Installed Traffic Signals

TAPAN K. DATTA

Installations of traffic signals at intersections are often associated with changes in the accident characteristics at those sites. Past studies have indicated a reduction in the severity of accidents, but in some instances accident frequencies have increased. Traffic signals also influence changes in the distribution of accident types. Head-on, left-turn accidents at signalized intersections are of great importance to traffic engineers. A study of newly installed traffic signals was performed in Michigan to evaluate the changes in the distribution of accident types, including head-on, left-turn accidents. This study involved statistical testing of before and after accident rates at groups of similar intersections. The results of this study indicated an increase in head-on, left-turn accidents at the group of signalized intersections with and without separate left-turn lanes. However, the rate of increase in such accidents at locations without a separate left-turn lane was not significant at the 95 percent confidence level. The group of locations in which a separate left-turn lane was installed along with the traffic signals also did not indicate a statistically significant change in the head-on, left-turn accident rate.

Traffic signals at intersections are often installed to alleviate traffic operational problems. However, such installations are often associated with changes in accident characteristics. Past studies at newly signalized intersections have concluded that the following:

- Accident frequency increases, as well as decreases, have been reported after traffic signal installation (*1*).
- Accident severity is generally reduced (*1*).
- Right-angle accident frequency decreases (*2–4*).
- Rear-end accidents generally increase (*2,4,5*).
- Head-on, left-turn accidents generally increase (*2*).
- Use of multiphase signals results in a reduction in head-on, left-turn accidents (*6*).

These generalized conclusions have been used by engineers to predict the impact on accident characteristics caused by the installation of traffic signals.

In October 1986 the Michigan Department of Transportation (*7*) sponsored a study of traffic signals at intersections scattered throughout the state. The signals were installed between 1978 and 1983. The objective of this study was to evaluate the change in accident characteristics at these newly signalized locations.

Department of Civil Engineering, Wayne State University, 5050 Anthony Wayne Drive, 2152 Engineering Building, Detroit, Mich. 48202.

The project included collection of all available data on the candidate traffic signal locations from existing files, plans, and diagrams. The data consisted of

- Intersection names, locations, and milepoints;
- Intersection geometry;
- Traffic volumes;
- Signal equipment;
- Signal timing and phasing; and
- Traffic accidents.

The study investigated the impact of traffic signal installations on accident characteristics. The results of the analysis on groups of intersections with and without left-turn lanes are presented, with particular emphasis on the head-on, left-turn accident rate.

DATA COLLECTION

A list of 235 candidate locations was initially selected as a part of this study. This time period was selected to provide 2 to 3 years of accident data for both the before and after conditions of the proposed study sites.

During the study a number of locations were eliminated from the candidate site list, either because they were of temporary nature due to construction activities or because they were not judged to be of interest (such as a traffic signal at a shopping center entrance). Also, several sites were eliminated because of a lack of traffic volume information. A total of 102 intersection locations were studied.

DATA ANALYSIS

The before and after accident rates were studied to evaluate the impact of signal installation. In this analysis the accident rate is defined as the number of accidents per million entering vehicles (accidents/MV).

The effect of signal installations was determined on the following accident types:

Accident Type	*Measure of Effectiveness*
Total accidents	Total accidents/MV
Right-angle accidents	Right-angle accidents/MV
Rear-end accidents	Rear-end accidents/MV
Injury accidents	Injury accidents/MV
Head-on, left-turn accidents	Head-on, left-turn accidents/MV
Other type accidents	Other type accidents/MV

Accident Rate Calculation Procedures

The procedure for calculating the accident rate consisted of three distinct efforts:

- Determination of before and after accident frequency from the statewide accident data base,
- Determination of before and after average daily traffic (ADT),
- Determination of accident rate.

Before and After Accident Frequency

The total number of accidents for the 3-year period before installation of the signal and the 3-year period after installation was identified. For signals installed after 1982 or before 1979, a 2- or 1-year study period was selected depending on the date of signal installation and availability of data.

Before and After ADT

The average traffic volume (total ADT of all approaches) for a 3-year period before and after the installation of the traffic signal at each location was calculated as follows:

Total ADT of all approaches = approach ADT on trunkline + approach ADT on minor streets

If only after or only before minor street ADT was available, the missing information was calculated using a growth factor (GF) as follows:

$$GF = \frac{\text{average ADT on trunkline after signal}}{\text{average ADT on trunkline before signal}}$$

Minor (before) street ADT = minor (after) street ADT/GF

Accident Rate

The accident rate for each location before and after signal installation was calculated using the following formula:

$$\text{Accident rate} = [A(10^6)]/(365TV)$$

where

A = accident frequency (total for the study period),
T = time period of the study (in years),
V = average daily traffic.

Test of Significance

The before and after accident rates for the groups of intersections were compared using paired t-tests to determine if there were statistically significant differences. The three groups of intersections tested were

- Groups with separate (exclusive) left-turn lanes,
- Groups without separate left-turn lanes, and
- Groups of intersections where the left-turn lane was added coincident to the signal installation.

The comparison of the t statistic with t critical was performed within each of these three groups for the before and after accident rates. The tests were two-tailed tests at a level of significance of .05.

EVALUATION OF HEAD-ON, LEFT-TURN ACCIDENT RATES

Head-on, left-turn accidents at intersections are of great concern to traffic engineers. Past studies have generally concluded that this type of accident increases after the installation of traffic signals. Intersections at which left-turn lanes existed and locations at which separate left-turn lanes did not exist were analyzed by comparing mean accident rates both before and after the installation. Locations at which a left-turn lane was installed coincident to signal installation were also analyzed. The results of this mean accident rate analysis are presented in Table 1 and shown graphically in Figure 1.

The statistical test of the mean rate of head-on, left-turn accidents showed significant differences at locations with left-turn lanes. In fact, the accident rate went up from 0.15 to 0.27. However, at locations without left-turn lanes and locations at which the left-turn lane was installed coincident to signal installation, the changes were not statistically significant.

ANALYSIS OF INTERSECTIONS WITH AND WITHOUT A LEFT-TURN LANE

As part of this analysis, comparisons of the before and after mean accident rates by total, rear-end, injury, head-on, left-

TABLE 1 RESULTS OF PAIRED t-TESTS FOR HEAD-ON, LEFT-TURN ACCIDENT RATE

Intersection Type	No. of Sites	Mean Accident Rate (Accidents/MV)		Percent (%) Increase (+) Decrease (-)	"t" Statistics	Two Tail t_c $\alpha = 0.05$
		Mean Before Accident Rate	Mean After Accident Rate			
Locations With Left-turn Lane	29	0.15	0.27	+80.0	-2.25*	2.048
Locations Without Left-turn Lane	35	0.12	0.19	+58.3	-1.58	2.030
Locations Where Left-turn Lane Was Installed Coincident to Signal Installation	14	0.27	0.29	+7.4	-0.22	2.160

* Means Significantly Different

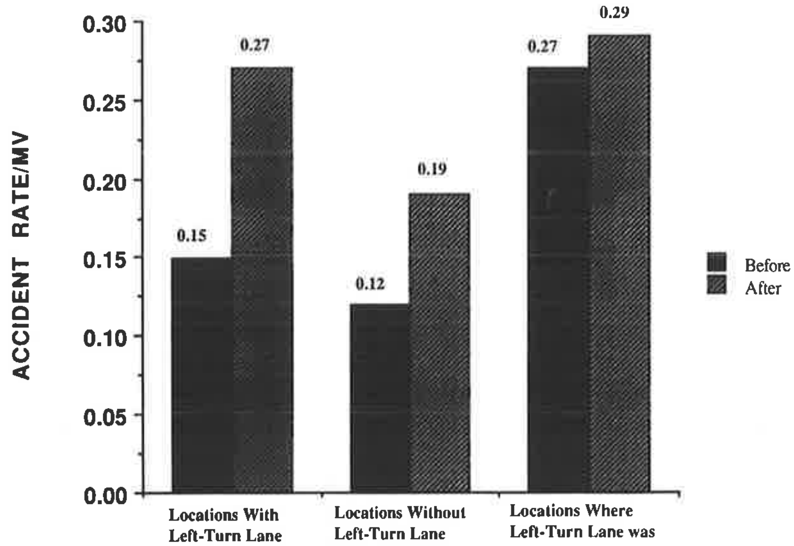


FIGURE 1 Before and after head-on, left-turn accident rates for locations with and without left-turn lane.

TABLE 2 RESULTS OF PAIRED t-TESTS FOR LOCATIONS WITH LEFT-TURN LANE

Accident Type	Mean Accident Rate (Accidents/MV) Mean Before Accident Rate	Mean After Accident Rate	Percent (%) Increase (+) Decrease (-)	"t" Statistics	Two Tail t_c $\alpha = 0.05$ d.f. 28
Total	1.98	1.78	-10.1	0.87	2.048
Right-angle	0.60	0.24	-60.0	4.06*	2.048
Injury	0.66	0.54	-18.2	1.34	2.048
Rear-end	0.32	0.57	+78.1	-2.46*	2.048
Head-on left-turn	0.15	0.27	+80.0	-2.25*	2.048
Other	0.91	0.70	-23.1	+2.08*	2.048

* Means Significantly Different

turn, and other types of accidents were performed. The results are presented in Tables 2–4 and shown graphically in Figures 2–4.

Table 2 presents the results of the paired t-tests for locations with left-turn lanes. For this group of 29 intersections, the following were observed:

- The total accident rate was reduced 10.1 percent, but the difference between before and after was not statistically significant.
- The right-angle accident rate was reduced by 60 percent, and the change was statistically significant.
- The injury accident rate was reduced by 18.2 percent, but the change was not statistically significant.
- The rear-end accident rate increased by 78.1 percent, which was statistically significant.
- The head-on, left-turn accident rate increased by 80 percent, which was statistically significant.
- The rate for other types of accidents was reduced by 23 percent, which was statistically significant.

Table 3 presents the results of the paired t-tests for locations without left-turn lanes. For this group of 35 intersections, the following were observed:

- The total accident rate was reduced by 23.6 percent, and this change was statistically significant.
- The right-angle accident rate was reduced by 51.7 percent, which was statistically significant.
- The injury accident rate was reduced by 11.8 percent, which was not statistically significant.
- The rear-end accident rate was reduced by 48.3 percent, which was statistically significant.
- The head-on, left-turn accident rate increased by 58.3 percent, but this change was not statistically significant.

TABLE 3 RESULTS OF PAIRED t-TESTS FOR LOCATIONS WITHOUT LEFT-TURN LANE

Accident Type	Mean Accident Rate (Accidents/MV) Mean Before Accident Rate	Mean After Accident Rate	Percent (%) Increase (+) Decrease (-)	"t" Statistics	Two Tail t_c $\alpha = 0.05$ d.f. 34
Total	1.74	1.33	-23.6	2.35*	2.03
Right-angle	0.58	0.28	-51.7	2.53*	2.03
Injury	0.51	0.45	-11.8	0.78	2.03
Rear-end	0.31	0.46	+48.3	-3.09*	2.03
Head-on left-turn	0.12	0.19	+58.3	-1.58	2.03
Other	0.74	0.41	-44.6	+4.58*	2.03

* Means Significantly Different

TABLE 4 RESULTS OF PAIRED *t*-TESTS FOR LOCATIONS WHERE LEFT-TURN LANE WAS ADDED COINCIDENT TO SIGNAL INSTALLATION

Accident Type	Mean Accident Rate (Accidents/MV)		Percent (%) Increase (+) Decrease (-)	"t" Statistics	Two Tail t_c $\alpha = 0.05$
	Mean Before Accident Rate	Mean After Accident Rate			
Total	2.08	1.42	-31.7	3.98*	2.16
Right-angle	0.74	0.27	-63.5	3.23*	2.16
Injury	0.77	0.43	-44.2	4.07*	2.16
Rear-end	0.31	0.39	+25.8	-1.59	2.16
Head-on left-turn	0.27	0.29	+7.4	-0.22	2.16
Other	0.76	0.47	-38.2	3.31*	2.16

* Means Significantly Different

- The rate for other types of accidents was reduced by 44.6 percent, which was statistically significant.

Table 4 presents the results of the paired *t*-tests for 14 locations at which the left-turn lane was constructed during the installation of the traffic signals. The following were observed from the analysis of this group of locations:

- The total accident rate was reduced by 31.7 percent, which was statistically significant.
- The right-angle and injury accident rates were reduced by 63.5 and 44.2 percent, respectively. Both of these changes were statistically significant.
- The rear-end and head-on, left-turn accident rates were increased by 25.8 and 7.4 percent, respectively. However, neither of these changes was statistically significant.
- The rate for other types of accidents was reduced by 38.2 percent, which was statistically significant.

FINDINGS

For intersections with and without a left-turn lane, the following were concluded:

- The mean head-on, left-turn accident rate was significantly different only for intersections with left-turn lanes in the before period.
- The mean head-on, left-turn accident rate at intersections with left-turn lanes increased 80 percent, compared with 58.3 percent for locations without left-turn lanes.

For intersections at which a left-turn lane was added coincident to signal installation, the following were concluded:

- The head-on, left-turn, accident rate increased by 7 percent.
- The total accident rate was reduced by 31.7 percent, which was statistically significant.
- Right-angle and injury accident rates were reduced by 63.5 and 44.2 percent, respectively, and the differences between the before and after accident rates were statistically significant.
- The rate for other types of accidents was reduced significantly (by 33.1 percent), as it was for locations with and without a left-turn lane.

The criteria for installation of traffic signals were assumed to be consistent with the warrants for traffic signals. No effort was made to determine if there were inconsistencies in signal installation criteria.

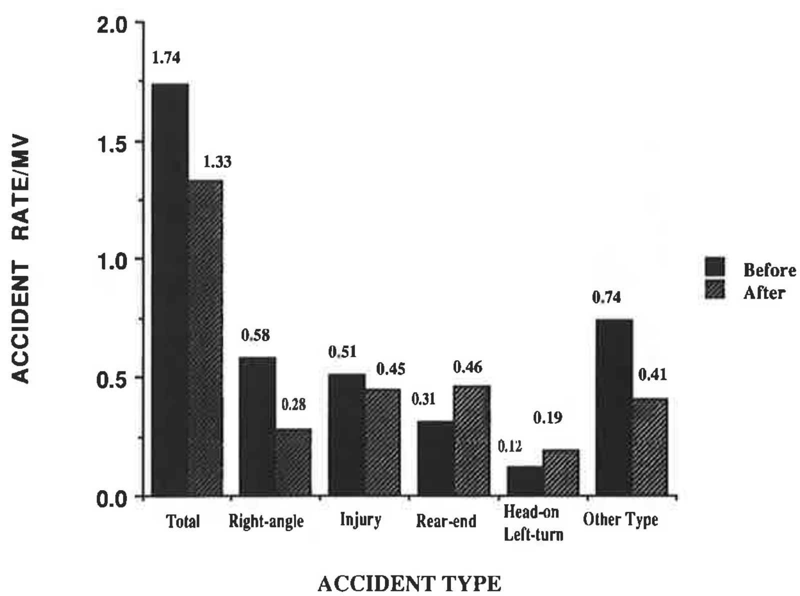


FIGURE 2 Before and after accident rates for locations without left-turn lane.

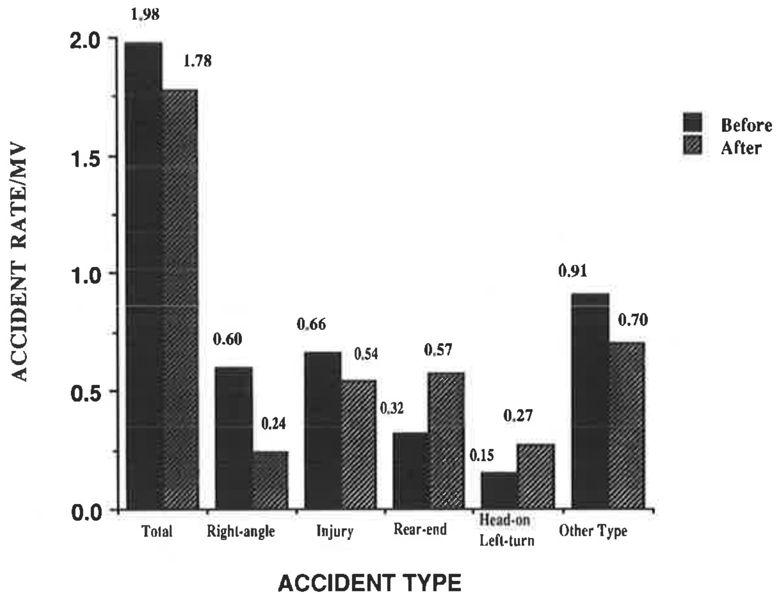


FIGURE 3 Before and after accident rates for locations with left-turn lane.

SUMMARY

The results of this study require additional commentary to make the findings usable for traffic and safety engineers. These comments are as follows:

- Locations with an exclusive left-turn indicated a statistically significant difference in head-on, left-turn accident rates before and after the signal installations. In fact, the study showed an 80 percent increase. This observation might be due to the following:
  - —Locations with an exclusive left-turn lane provide a better opportunity for left turns, thus attracting more left-turning traffic and creating increased accident potential.
  - —Intersections with exclusive left-turn lanes may often require a left-turn phase. In the state of Michigan, the left-turn phase is rarely included when a traffic signal is initially installed. Many of the intersections might have been safer with a separate left-turn phase.
  - —The absence of specific warrants for installation of a left-turn phase may have contributed to the absence of this phase in many locations where it could have reduced head-on, left-turn accidents.
- The significant decrease in right-angle accidents in all cases verifies past research results.
- In most categories, the increase in rear-end accident rates also verifies historical observations.
- The decrease in total accident rates at locations without left-turn lanes and at locations at which a left-turn lane was

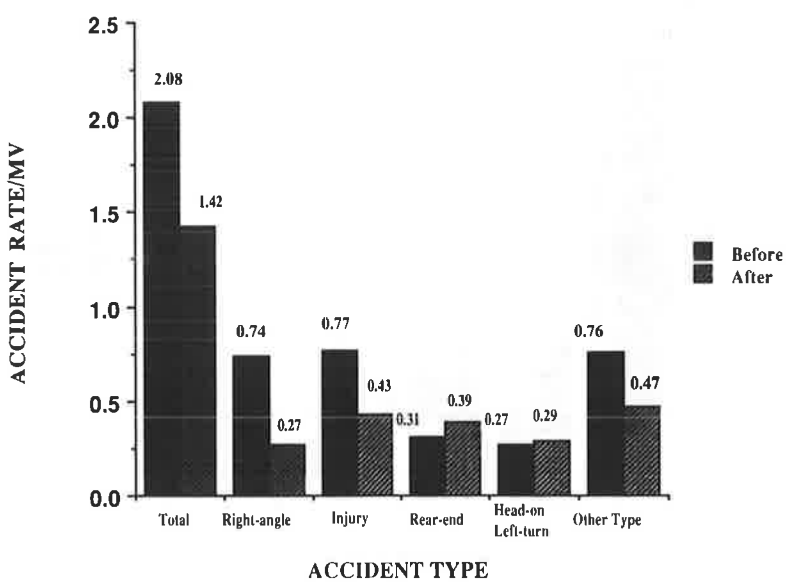


FIGURE 4 Before and after accident rates for locations where left-turn lane was added coincident to signal installation.

added coincident to signalization was somewhat unique and should be studied with a larger sample of intersection sites.

REFERENCES

1. D. Solomon. Traffic Signals and Accidents in Michigan. *Public Roads*, Vol. 30, Oct. 1959, pp. 234–237.
2. M. N. Clyde. Michigan Study Indicates Signals Increase Accidents. *Traffic Engineering*, Vol. 35, Nov. 1964, p. 32.
3. A. F. Malo. Signal Modernization. In *Special Report 93: Improved Street Utilization Through Traffic Engineering*, HRB, National Research Council, Washington, D.C., 1967, pp. 96–126.
4. G. F. King and R. B. Goldblatt. Relationships of Accident Patterns to Type of Intersection Control. In *Transportation Research Record 540*, TRB, National Research Council, Washington, D.C., 1976.
5. J. T. Hanna, T. E. Flynn, and W. K. Tyler. Characteristics of Intersection Accidents in Rural Municipalities. In *Transportation Research Record 601*, TRB, National Research Council, Washington, D.C., 1976.
6. K. R. Agent and R. C. Deen. Warrants for Left-Turn Signal Phasing. In *Transportation Research Record 737*, TRB, National Research Council, Washington, D.C., 1979.
7. T. K. Datta and U. Dutta. *Traffic Signal Installation and Safety*. Report FHWA-MI-RD-88-04. Michigan Department of Transportation; FHWA, U.S. Department of Transportation, Feb. 1989.

Publication of this paper sponsored by Committee on Traffic Records and Accident Analysis.

Influence of Road Width on Accident Rates by Traffic Volume

RACHEL GOLDSTINE

A study was conducted to assess the effect of shoulder and road widening on accident rates. Twenty-five projects covering 152 mi of road were selected for analysis. Sampled roads had been widened to one of four widths: 32, 36, 40, or 44 ft. Accident rates were compared before and after the construction period. Reductions of 38 to 53 percent were observed in accident rates, although the amount of reduction varied with traffic volume and the roadway width after construction. Accident rate reduction for the sampled projects was statistically significant at the 95 percent confidence level for before and after comparisons on most of the roads. Further research of the effects of resurfacing, restoration, and rehabilitation projects on accident rates in New Mexico is recommended.

Two-lane rural highway safety has a high priority in transportation research because more fatal accidents occur on these roads than on other roads. New Mexico has 9,680 mi of two-lane rural highways, which comprise 83 percent of all its state highways. Twenty-eight percent of the total million vehicle-miles traveled (MVMT) are on two-lane rural roads.

The majority of resurfacing, restoration, and rehabilitation (3R) projects consist of widening, resurfacing, and overlaying existing roads. In addition, protective barriers may be built to shield vehicles from roadside objects, or these objects may be removed. Such measures as culverts and drain facilities to prevent water flow from washing out roads are other possible improvements undertaken in 3R projects. Also, horizontal and vertical curves may be realigned, although the realignment is usually minor (major realigning of roads is undertaken in 3R reconstruction projects).

Traffic volume has been identified by Cleveland et al. (*1*) to be the most important factor in the frequency of accidents for two-lane rural roads with annual average daily traffic (AADT) below 3,000. Another factor that influences accidents is clear recovery zones (*2,3*). Some researchers (*1,4*) have concluded that intersections have a stronger influence than other geometric characteristics on accident frequency.

The effects of a few dominant elements, such as average horizontal curvature (sum of external horizontal angles between tangents divided by segment length) and average vertical curvature (sum of vertical distances between crest and following sag divided by segment length), have been noted (*1,4–7*). Gradient has been found to have a weak effect when it is not accompanied by other geometric characteristics, such as intersections or horizontal curves (*4,6,8*).

New Mexico State Highway and Transportation Department, 1120 Cerrillos Road, South Building 1, Santa Fe, N.Mex. 87501.

It is difficult to identify the contribution of geometric elements to accident frequency because the design elements of existing roads interact (*1*). The findings of the studies that were reviewed were, in some cases, conflicting.

In particular, the relationships among road width, accident rate, and accident severity have not been established. Some studies have claimed small or even negative effects on the accident rate when roads are widened (*1,4,6,8*). Most studies, however, indicate that widening roads reduces accident rates, although at least one study has claimed that the reduction in accidents occurs with shoulder widening up to 8 ft after which there is an increase in accident rates.

A study conducted by the California Department of Transportation (*9*) investigated the relationship between shoulder widening projects and accident rates. The conclusion was that the safest widths depend on the volume of traffic. Several specific pairings between AADT and road width were supported: (a) if AADT is less than 3,000, a road should be widened to at least 32 ft (b) if AADT is between 3,000 and 5,000, the road should be widened to 32 ft; and (c) if AADT is more than 5,000, the road should be widened to 40 ft. These recommended widths were based on the following findings:

AADT	*Pavement Width after Widening (ft)*	*Percent Reduction in Accidents*	*Significant?*
<3,000	28	16	No
3,000–5,000	32	35	Yes
>5,000	40	29	Yes

It was concluded that there was no point in widening a 24-ft road to 28 ft because the reduction in accidents would not be great enough.

In TRB *Special Report 214* (*7*), the recommended minimum widths on highways with running speeds over 50 mph and more than 10 percent trucks are as follows:

AADT	*Pavement Width (ft)*
<750	24
751–2,000	30
>2,000	36

The criteria for widening roads in New Mexico are functional classification, AADT, and design speed. Almost all of the AADT values for the sampled roads are below 3,000. In general, roads with lower AADT have smaller widths than those with higher AADT. The current study verifies the relationships among accident rates, AADT, and road widening that were found in the California study (*9*) and the TRB special report (*7*).

OBJECTIVES

There were two objectives for this study. First, road widening and accident rates were matched for rural two-lane highways. Second, the findings of two studies (*7,9*) that AADT, road widening, and accident rates are related were confirmed.

DATA COLLECTION

Data were collected on two-lane rural roads classified as either federal-aid primary (FAP) roads or federal-aid secondary (FAS) roads. FAP and FAS roads were chosen because the majority of 3R projects are constructed on these roads. All projects completed from early 1981 to late 1986 were collected. Projects were rejected if they met one or more of the following four conditions:

1. Projects costing less than $200,000 per mile,
2. Projects that included major intersections,
3. Projects in urban areas, and
4. Projects having more than two lanes.

Twenty-five projects were chosen, comprising 152 mi of roadway. Project length varied from 1.5 to 15 mi.

Data collected for each project included project identification (ID), route, beginning and ending milelog, and estimated AADT. The projects were located using a milepoint system used by the New Mexico State Highway and Transportation Department (NMSHTD) in their consolidated highway data base. These points were used to match the appropriate AADT values to the segments and to match accidents to the appropriate segments.

Geometric information was collected from the as-built plans, including information on road widths, project lengths, sideslopes, average vertical curvature, and average horizontal curvature. Construction dates for projects were obtained from the districts that oversaw construction.

AADT values were obtained from a log used to record road segments and AADT. Most of the AADT values came from portable counters and are short-term counts. The counts were factored to an annual statistic using a growth statistic from permanent counters. Before 1988, the short-term counts were taken only once every 4 to 5 years; in between, the counts were factored to AADT using growth factors. These growth factors often did not match the actual rate of growth. Therefore, straight-line estimation was used rather than the growth factors for the years between the counts. The AADT varied from 300 to 3,700 for all surveyed projects.

Table 1 presents the pavement widths, number of projects, total project miles, and AADT (rounded to the nearest hundred) for projects included in this study.

Accident records were obtained from computer files kept by NMSHTD for a period beginning in 1978 and ending in 1988. The accident files are from reports filed by the New Mexico state police. The following variables were obtained from the accident files: accident location, date of accident, severity of accident (fatal, injury, or property damage), number of people killed or injured, and accident classification (e.g., fixed-object, other-vehicle, overturn, and parked-vehicle). The accident location was given in terms of milepoint and was reported to be accurate within 1/10 mi.

The years of accidents that could be included were limited because 1988 was the latest date that accidents had been recorded and accident records before 1978 were not available on the computer. Also, the period of time before construction and the period after construction had to be matched for the available accident records. One project had a 1½-year period for which accident records could be used for the before-construction period and for the after-construction period. All other projects were matched with at least 2 years of accident records before and after construction.

Accident rates per million vehicle-miles (AMVM) for the before and after periods were calculated using the average AADT values from each period. The following formula was used to compute the rates:

$$\text{AMVM} = \frac{\text{Number of accidents} \times 1{,}000{,}000}{\text{AADT} \times \text{segment length} \times \text{number of years} \times 365} \tag{1}$$

MVMT was computed using the following formula:

$$\text{MVMT} = \frac{\text{AADT} \times \text{segment length} \times \text{number of years} \times 365}{1{,}000{,}000} \tag{2}$$

DATA ANALYSIS

T-tests were performed on the data using the SAS Institute statistics package. In a *t*-test, a rate is first found for individual projects and then the rates for individual projects are averaged together by widening group (e.g., one group would be all roads widened to 32 ft), instead of computing one before and one after rate for a whole group. The *t*-test was based on the rate averages presented in Table 2.

Two different *t*-tests at the 95 percent level were used with the data. One of the *t*-tests was a comparison of group means *t*-test. The other *t*-test, a paired comparisons *t*-test, was used with $\log_{10}$-transformed data.

The paired comparisons *t*-test was used to measure before and after accident rates more precisely because, in this procedure, accident rates are contrasted before and after construction for the same road segment. Data used in the paired comparisons *t*-test were normalized by a $\log_{10}$ transformation.

TABLE 1 PROJECT DATA

Roadway Width After	Roadway Width Before	Number of projects	Total Project Miles	Annual Average Daily Traffic (AADT)
32	26–32	8	70.48	1000
36	26–28	5	18.60	1000
40	24–40	7	36.89	1500
44	24–32	5	26.56	2200
Total		25	152.53	

TABLE 2 RATE AVERAGES USED IN *t*-TEST

After Construction Width	Rate Before Construction	Rate After Construction	Percent Reduction
32 feet	2.70	.79	71%
36 feet	2.15	1.61	25%
40 feet	1.37	.89	35%
44 feet	1.24	.59	52%

The results for the *t*-test for group means and for the paired comparisons *t*-test with transformed data were the same.

No significant differences in accident and severity rates were found between the before and after construction periods for groups of projects widened to 36 and 40 ft. For road groups widened to 32 and 44 ft, the test results were significant.

Because the group results were inconsistent with some of the individual project accident rate reductions, the data were segregated into various AADT groupings. Table 3 gives some indication of the road widths necessary for safety benefits by AADT. The available projects are small in number for each cell, but they form a pattern in the degree of road widening, AADT, and accident rate changes.

Table 3 indicates that there is no safety benefit if a road is widened to 32 ft and AADT exceeds 1,000. Four projects that were widened from 28 to 32 ft and had AADT in the range of 1,001 to 2,000 had no significant reduction in accidents. The property-damage-only (PDO) accident rate actually increased on these projects, whereas overturn rates remained constant for the before and after periods.

Four projects with the range of 0 to 1,000 AADT did have a statistically significant reduction in accidents. Two of these projects were 26 ft wide, and the other two were 32 ft wide. Therefore, widening rural two-lane roads from 26 to 32 ft, with a combination of other improvements allowed in 3R practices, may reduce the number of accidents when AADT is under 1,000.

Accident severity and overturn rates decreased sharply for the three projects that were widened to 36 ft and had AADT between 1,000 and 2,000.

Five roads with AADT values less than 2,000 were widened to 40 ft. There was no significant reduction in accidents for these projects. Two projects in the group with more than 3,000 AADT exhibited significant accident reduction although the width of the road was not changed.

Four projects widened to 44 ft and having AADT > 2,000 exhibited reductions in accidents; however, a 40-ft width appears to be adequate for even the highest category of AADT. Further study will be required to more firmly establish the widths needed for the greatest safety.

REVISED DATA

On the basis of the findings presented in table 3, a road should be widened to 32 ft if AADT < 1,000, to 36 ft if it has AADT between 1,000 and 2,000, and to 40 or 44 ft if AADT > 2,000. Table 4 presents a summary of the reduction in percent.

ACCIDENT TREND OVER TIME

The time period included in this study is more than a decade. During this period accident rates on rural two-lane highways

TABLE 3 TOTAL ACCIDENTS BY AADT GROUPS

Width After	AADT	0 - 1000 Acc Rate	1001 - 2000 Acc Rate	2001 - 3000 Acc Rate	Over3000 Acc Rate
32	# Proj	(4)	(4)		
	Before	28 3.94	63 1.02		
	After	4 0.48	52 0.97		
	% Reduc	88*	5		
36	# Proj	(2)	(3)		
	Before	3 1.48	43 2.57		
	After	4 1.60	29 1.35		
	% Reduc	-8	47*		
40	# Proj	(3)	(2)		(2)
	Before	8 1.80	23 1.66		55 0.92
	After	4 0.49	20 0.91		31 0.60
	% Reduc	73	45		35*
44	# Proj	(1)		(3)	(1)
	Before	3 1.00		59 1.26	27 1.46
	After	3 0.98		25 0.47	9 0.58
	% Reduc	0		63*	60*

* Statistically Significant Reduction

TABLE 4 REVISED ACCIDENT RATE REDUCTION IN PERCENT

Roadway Width After	AADT	Total	Injury	PDO	Overturn
32	<1000	88	76	92	95
36	<2000	47	72	31	83
40	>2000	35	-7	52	10
44	>2000	63	47	71	63

changed because of new traffic laws and better vehicles. To account for these and other changes, the accident rates were adjusted for the accident trend on the routes as a whole.

The accident trends for the routes on which the projects took place were computed using a series of steps. Nineteen routes were identified for the 25 projects. The annual accident rates for the entire routes, with the exception of urban routes, were computed. Then, these annual accident rates were averaged within each widening group. The before and after periods for the projects were matched with the rates computed for the corresponding routes. As shown in Figure 1, not all years of actual data were available for the routes as a whole. When data were missing, the existing data were weighted more heavily. A percent reduction was determined for each individual project by comparing the average rates of before and after periods, using Figure 1. These reductions were averaged within each widening group, as follows:

Roadway Width After (ft)	*Percent Reduction*
32	35
36	2
40	32
44	25

The accident rates were reduced by 25 to 35 percent for all road groups except for roads widened to 36 ft. These roads had only a 2 percent reduction in their accident rates.

In the following table the corrected change in percent from before construction to after construction is shown for accident rates:

Roadway Width After (ft)	*AADT*	*Percent Reduction*
32	<1,000	53
36	<2,000	45
40	>2,000	3
44	>2,000	38

The corrected percentages are the result of subtracting the percent reduction in accident rates attributed to time trends

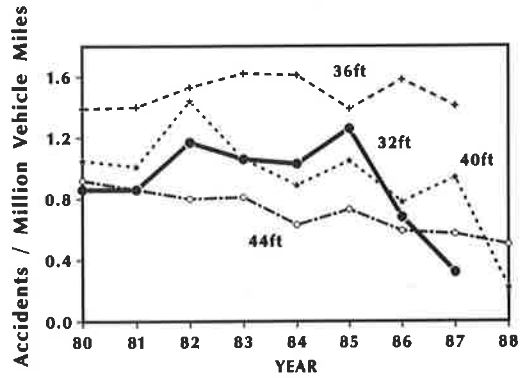


FIGURE 1 Trends of accident rates.

from the percent reduction in accident rates calculated from the project data. Most of the corrected percentages are still high. The only exception is the corrected accident rate percentage change of 3 percent for 40-ft roads with AADT values above 2,000. Because the two projects were 40 ft before construction, the small reduction in accidents is not surprising.

The corrected accident rate reductions may be due partly to factors other than road widening, such as improved pavement structure, improved riding quality, improved skid resistance of pavement surface, protection from roadside obstacles, improved signing, striping, and minor curve improvement.

ACCIDENT CLASSIFICATION

Accidents are categorized for roads widened to 32, 36, 40, and 44 ft in Tables 5 and 6. Most accidents fell within one of the four categories: overturn, other-vehicle, fixed-object, and animal. Table 7 indicates that the number of accidents was reduced by 45 percent for the overturn and other-vehicle class, 24 percent for the fixed-object class, and 33 percent for the animal class.

PARKED VEHICLES

The influence of road widening on the number of accidents classified as a parked-vehicle accident was also investigated, although this class is rare in the accident records. Table 5 indicates that only 2 percent of all surveyed accidents were related to parked vehicles. No further analyses were conducted.

TABLE 5 ACCIDENT CLASSIFICATION BY PERCENT

CLASS	FREQUENCY	PERCENT	CUMULATIVE FREQUENCY	CUMULATIVE PERCENT
ANIMAL	88	14.8	88	14.8
FIXED-OBJECT	93	15.6	181	30.4
OTHER-NON-COLLISION	33	5.5	214	35.9
OTHER-OBJECT	4	0.7	218	36.6
OTHER-VEHICLE	172	28.9	390	65.4
OVERTURN	191	32.0	581	97.5
PARKED-VEHICLE	12	2.0	593	99.5
PEDALCYCLIST	2	0.3	595	99.8
PEDESTRIAN	1	0.2	596	100.0

TABLE 6 SEVERITY BY TIME GROUP

FREQUENCY / PERCENT / ROW PCT / COL PCT	AFTER	BEFORE	CONSTRUCTION	TOTAL
FAT	10 1.68 45.45 5.52	8 1.34 36.36 2.56	4 0.67 18.18 3.88	22 3.69
INJ	65 10.91 30.95 35.91	112 18.79 53.33 35.90	33 5.54 15.71 32.04	210 35.23
PDO	106 17.79 29.12 58.56	192 32.21 52.75 61.54	66 11.07 18.13 64.08	364 61.07
TOTAL	181 30.37	312 52.35	103 17.28	596 100.00

CONCLUSIONS

The 3R improvements on surveyed rural two-lane FAP and FAS roads in New Mexico significantly reduced the accident rates for some of the AADT groups when the roads were widened to 32, 36, 40, and 44 ft.

The accident rate was reduced by 53 percent for roads widened to 32 ft when AADT was less than 1,000. The accident rate was reduced by 45 percent for roads widened to 36 ft when AADT was between 1,000 and 2,000. Determining the performance of roads 36 ft wide with AADT > 2,000 will require further investigation. The accident rate was reduced by 38 percent for roads widened to 44 ft with AADT > 2,000.

The results of the study support the practice of road widening to reduce accidents for two-lane rural roads. Also, the findings are consistent with the relationships found in the California study (*9*) and in the minimum standards set in TRB *Special Report 214* (*7*). In other words, the higher the AADT, the wider the road should be.

DISCUSSION OF RESULTS

The study results support the general findings of the California study (*9*) and TRB Special Report 214. (*7*). However, 3R practice in New Mexico requires a wider minimum finished roadway. This study indicates that the greater minimum width is justified by reduced accident rates.

Initially, there were some problems identifying which projects were 3R projects. Projects in the NMSHTD system are coded according to work-type descriptors. These descriptors, however, are not always accurate. Eventually, the basis for choosing the projects were the dollar amount per mile because this factor was independent of the coding.

Two other problems that could not be overcome in this research limited the analyses and strength of the conclusions. One was the small number of projects available for analysis, and the other was that the work on the projects was not

TABLE 7 CLASSIFICATION BY TIME GROUP

FREQUENCY / ROW PCT	AFTER	BEFORE	CONSTRUCTION	TOTAL
ANIMAL	30 34.09	45 51.14	13 14.77	88
FIXED-OBJECT	32 34.41	42 45.16	19 20.43	93
OTHER-NON-COLLISION	8 24.24	18 54.55	7 21.21	33
OTHER-OBJECT	0 0.00	3 75.00	1 25.00	4
OTHER-VEHICLE	47 27.33	85 49.42	40 23.26	172
OVERTURN	60 31.41	110 57.59	21 10.99	191
PARKED-VEHICLE	3 25.00	7 58.33	2 16.67	12
PEDALCYCLIST	0 0.00	2 100.00	0 0.00	2
PEDESTRIAN	1 100.00	0 0.00	0 0.00	1
TOTAL	181	312	103	596

uniform. The 25 projects included in the research had various structural changes: 52 percent had changes in their sideslopes, 16 percent had significant changes in vertical curvature, and 8 percent had changes in both the sideslopes and vertical curvature. Eight, or 32 percent, of the projects had no significant changes in either their sideslopes or their vertical curvature. This variety in the projects means that, for any detailed analysis of the projects' geometric characteristics, the individual groupings would have to be very small.

Inferences drawn from accident rates computed for two-lane rural road segments should be tentatively drawn because these roads tend to have more variable accident rates than those for rural roads as a whole. This is partly true because a low AADT in the equation for accident rate means that each accident has a greater effect on the calculated rate. Also, the absolute number of accidents will be smaller on low-volume roads, so each accident will have a greater effect.

The combined effects of the small sample size, variety in projects, and variable accident rates on two-lane rural roads mean that the researcher must look at data over a greater time period for these roads than for other road classifications.

The current study will be continued using additional data from 3R projects and the corresponding accident data. A more precise identification of the effects of the road improvements on accidents can then be drawn.

ACKNOWLEDGMENTS

Several NMSHTD members assisted in this research. Parviz Eftekhari chose the topic of 3R projects and accidents. He also directed and participated in much of the research. Olga Pendelton of the Texas Transportation Institute assisted in choosing the statistical analyses for this research. Jerry Hall of the University of New Mexico gave practical advice and criticism. Richard Boyce contributed Figure 1 to this paper.

REFERENCES

1. D. E. Cleveland, L. P. Kostyniuk, and K.-L. Ting. Design and Safety on Moderate-Volume Two-Lane Roads. In *Transportation Research Record 1026*, TRB, National Research Council, Washington, D.C., 1985, pp. 51–61.
2. J. L. Graham and D. W. Harwood. *NCHRP Report 247: Effectiveness of Clear Recovery Zone*. TRB, National Research Council, Washington, D.C., 1982.
3. C. V. Zegeer, D. W. Reinfurt, W. W. Hunter, J. Hummer, R. Stewart, and L. Herf. Accident Effects of Sideslope and Other Roadside Features on Two-Lane Roads. In *Transportation Research Record 1195*, TRB, National Research Council, Washington, D.C., 1988, pp. 33–47.
4. J. K. Kihlberg, and K. J. Tharp. *NCHRP Report 47: Accident Rates as Related to Design Elements of Rural Highways*. HRB, National Research Council, Washington, D.C., 1968.
5. M. S. Raff. *Interstate Highway Accident Rates on Federal-Aid and Other Highway Systems, 1976*. FHWA, U.S. Department of Transportation, 1978.
6. P. A. McBean. *The Influence of Road Geometry at a Sample of Accident Sites*. TRRL Report 1053. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, 1982.
7. *Special Report 214: Designing Safer Roads. Practices for Resurfacing, Restoration and Rehabilitation*. TRB, National Research Council, Washington, D.C., 1987.
8. *NCHRP Report 197: Cost and Safety Effectiveness of Highway Design Elements*. TRB, National Research Council, Washington, D.C., 1978.
9. E. A. Rinde. *Accident Rate vs. Shoulder Widening*. California Department of Transportation, Sacramento, 1977.

Publication of this paper sponsored by Committee on Traffic Records and Accident Analysis.

Identification of Accident-Prone Locations in Greater Amman

FOUAD A. GHARAYBEH

A study was conducted to identify and rank dangerous traffic locations in Greater Amman, the capital of Jordan. Accident frequency, accident rate, critical rate, and accident seriousness measures were incorporated to develop a technique for ranking these sites for future improvement. It was proven that the use of accident frequency as the sole criterion for identifying or ranking hazardous locations may be misleading.

Jordan is a Middle Eastern country. It has about 3 million inhabitants in an area of 96 000 km². Traffic accidents are a major cause of loss of life and wealth in Jordan (*1*). In 1988 the country had about 242,000 registered vehicles and 5 500 km of paved roads (*2*). From 1983 through 1988 the country experienced an average of 16,000 accidents, 9,000 injuries, and 424 deaths per year (*3*).

Amman, the capital of Jordan, has been facing rapid population growth and migration from rural areas. It has about 35 percent of the country's population (1 million) and 82 percent of its registered vehicles. Traffic accidents in the city count for 53 percent of those in Jordan as a whole (*3*).

No sound traffic safety program has been established in Amman. All safety measures are taken by the police department on the basis of experience and intuition, with little knowledge of basic traffic accident analyses procedures and the essentials of traffic safety programs implementation.

Because of these concerns and others mentioned in the literature (*1–4*), and also to establish a base for further traffic safety research, a study was carried out as part of a comprehensive traffic safety project. The project aims to improve traffic safety in Amman through enhancing traffic control systems, implementing safety countermeasures, and distinguishing hazardous locations. The objective of this study was to identify accident-prone locations in Greater Amman and to establish a ranking system for these locations for future improvement funding.

DATA ACQUISITION

Accident records were obtained from police department files. Any site that had 10 or more accidents per year was selected. On that basis, 28 intersections and nine roads were found to need further investigation. The naming and numbering system of Greater Amman streets is under development. Some of the sites still have common names. Following is a full description of the study locations:

- *Site 1.* Central vegetable supermarket entrance;
- *Site 2.* Military hospital entrance from King Abdullah Street, Marka;
- *Site 3.* Elled Society intersection: intersection of Nablus and Al-Jalil streets, Jabal Al-Hussien;
- *Site 4.* Custom Triangle: Madaba and Sahab roads intersection near Al-Jumrock;
- *Site 5.* Al-Basheer Hospital entrance from Al-Taj Street, Al-Ashrafieh;
- *Site 6.* Al-Marbat Bridge: end of Estiklal Street at Al-Mahattah;
- *Site 7.* Driving and vehicle license department entrance from King Abdullah Street, Marka;
- *Site 8.* Television Triangle;
- *Site 9.* Ain Gazal intersection: end of Ain Gazal Road at Al-Jaish Street (intersection of Al-Jaish and Biet Ummer streets);
- *Site 10.* Marka Shneller Bridge;
- *Site 11.* Palace Mountain entrance from Wadi Al-Haddadeh (King Hussien Ben Ali Street);
- *Site 12.* Middle East intersection: intersection of Yarmouk and Prince Hassan streets;
- *Site 13.* Abdoun Triangle, Wadi Abdoun;
- *Site 14.* Al Nasha intersection: intersection of King Abdullah, Yarmouk, and Al-Estiklal streets;
- *Site 15.* West Broadcasting Triangle at Nauor Road;
- *Site 16.* Qal'ah intersection: intersection of Khalid Ben Al-Walid and King Hussein Ben Ali streets;
- *Site 17.* Al-Taj intersection: end of Al-Taj Street at King Abdullah Street;
- *Site 18.* Al-Hussien Medical Center entrance;
- *Site 19.* Samir Al-Refa'i School intersection: intersection of Prince Mohammed and Al-Ma'amoon streets;
- *Site 20.* First Circle, Jabal Amman (roundabout);
- *Site 21.* Abdoun intersection near Abdoun supermarket;
- *Site 22.* Seventh Circle, Al-Sowaifieh (signalized intersection);
- *Site 23.* Second Circle, Jabal Amman (roundabout);
- *Site 24.* Fifth Circle, Jabal Amman (signalized intersection);
- *Site 25.* Sixth Circle, Um Odeinah (signalized intersection);
- *Site 26.* Wadi Sakrah intersection: intersection of Al-Sharif Naser Ben Jamil, Al. Sharif Shaker Ben Zaid, Kendy, and Ben Sinai streets;
- *Site 27.* Third Circle, Jabal Amman (roundabout);

Faculty of Engineering, Mu'tah University, P.O. Box 7, Mu'tah, Karak, Jordan.

- *Site 28.* Fourth Circle, Jabal Amman (signalized intersection);
- *Site 29.* Mesdar and Prince Hassan streets;
- *Site 30.* Thirtieth and Osama Ben Zaid streets;
- *Site 31.* King Abdullah Street, Marka;
- *Site 32.* Bayader Wadi-Essair Main Street;
- *Site 33.* Kalid Ben Al-Walid Street, Jabal Al-Hussien;
- *Site 34.* Al-Estiklal Street;
- *Site 35.* Yarmouk Street;
- *Site 36.* Gardens and Wasfi At-Tel streets, Tla' Al-Ali; and
- *Site 37.* University and Queen Alia roads west of University Bridge.

Traffic counts were carried out on the selected sites during the summer season (July to October) and were extrapolated for the rest of the year using expansion factors. The expansion factors were determined using average monthly transportation petrol sales in the city, obtained from Jordan Petroleum Refinery Company (*5*). Intersection data were sorted into two groups depending on traffic stream characteristics, such as daily variation in traffic volume, trip purpose, geographic location, percentage of heavy vehicles, and type of intersection control.

METHODS FOR IDENTIFICATION OF ACCIDENT-PRONE LOCATIONS

A literature review revealed that there was no specified method for verification of accident-prone locations. However, methods for identification of hazardous locations have been investigated by many researchers.

Maher and Mountain (*6*) adjudged the method of ranking sites according to their annual accident total (AAT) as being simple and attractive. Plass and Berg (*7*) compared and evaluated conventional and opportunity-based accident rate expressions. Chira-Chavala and Mak (*8*) developed an algorithm to identify accident factors that are overrepresented at a site. Higle and Witkowski (*9*) developed a Bayesian model for identification of hazardous locations. Al-Esa et al. (*10*) measured the degree of hazard at urban intersections by the number of traffic conflicts. Lin (*11*) reviewed different methods used for identifying accident-prone locations and developed a model using a ranking system that combines the results of four classical methods. A brief description of these methods follows.

Accident Frequency Method

In the accident frequency method the number of accidents within a specific time period is used to determine the priority sequence for safety or improvement funding.

Accident Rate Method

The accident rate method projects the accident rate on the basis of exposure data, such as the traffic volume or the length of the road section being considered.

Accident Possibility Method

The accident possibility method is also called the quality control method. It is based on the critical accident rate (RC), as shown in the following equation:

$$\mathrm{RC} = \mathrm{Ra} + K(\mathrm{Ra}/M)^{0.5} + 1/2M \qquad (1)$$

where

- Ra = average accident rate for the group that contains the site under study,
- K = statistical rate factor with specified significance level (for confidence level = 95 percent, K = 1.645), and
- M = exposure in million entering vehicles (MEV) or 100 million vehicle-km (for road sections).

Accident Seriousness Method

The accident seriousness method is used to compare accidents at different locations by assigning weights to each accident depending on its severity. The equivalent total accident number (ETAN) is calculated for all sites and used for the ranking system.

DATA ANALYSIS

Accidents recorded in 1988 were analyzed. Accident data for the previous year were also available, but because significant intersection improvement plans had been undertaken during that year, analogous criteria between the years could not be achieved.

Calculations of accident rates in accordance with the four methods previously discussed were incorporated to find a combined method that could be used to identify accident-prone locations.

Calculations in agreement with the first three methods are presented in Table 1. The table shows the accident frequency, accident rate (R), critical rate (RC), and danger factor (DF) for the sites under study. The accident rate is the accident frequency divided by the exposure. The critical rate was calculated using Equation 1 with a confidence level equal to 95 percent (K = 1.645). The danger factor is the accident rate divided by the critical rate (DF = R/RC).

Table 1 indicates that sites with a higher accident frequency are not necessarily those with higher accident rates or a higher danger factor. This finding goes against the prevailing judgment used by personnel of the police department, who consider sites with high accident frequency as more dangerous. (This phenomenon is addressed later in Table 3.) On the other hand, no valid relationship could be found between traffic volume and the accident frequency, accident rate, or danger factor, which indicates that these measures are not necessarily proportional to traffic volume. The table also indicates that Sites 1–5 (unsignalized intersections) and Sites 29–32 are accident-prone locations according to the quality control method criteria. These sites have a danger factor greater than 1 and need imperative modifications. On the other hand, none of the Group 2 sites (signalized intersections) are in this cate-

TABLE 1 TRAFFIC VOLUME AND ACCIDENT RATES

SITE NO.	LENGTH (KM)	AADT (VEH.)	EXPOSURE (MIL.VEH.)	NO.OF ACC.	RATE (R)	CRITICAL RATE(RC)	DANGER FACTOR(DF)
GROUP 1							
1		18964	6.922	33	4.77	3.25	1.47
2		21612	7.888	33	4.18	3.15	1.33
3		14434	5.268	24	4.56	3.47	1.31
4		29083	10.615	34	3.20	2.97	1.08
5		15192	5.545	20	3.61	3.42	1.05
6		51845	18.923	50	2.64	2.71	0.98
7		27885	10.178	18	1.77	3.00	0.59
8		34734	12.678	20	1.58	2.88	0.55
9		40053	14.619	22	1.50	2.81	0.53
10		32960	12.030	16	1.33	2.91	0.46
11		23310	8.508	12	1.41	3.10	0.45
12		58320	21.287	25	1.17	2.67	0.44
13		26955	9.839	13	1.32	3.02	0.44
14		50916	18.584	20	1.08	2.72	0.40
15		19302	7.045	10	1.42	3.23	0.44
16		33294	12.152	12	0.99	2.90	0.34
17		37872	13.823	10	0.72	2.84	0.25
GROUP 2							
18		27792	10.144	29	2.86	3.00	0.95
19		35513	12.962	25	1.93	2.87	0.67
20		27989	10.216	16	1.57	2.99	0.52
21		29724	10.849	16	1.47	2.96	0.50
22		58793	21.459	27	1.26	2.66	0.47
23		42927	15.668	17	1.08	2.78	0.39
24		66361	24.222	23	0.95	2.62	0.36
25		58721	21.433	20	0.93	2.66	0.35
26		58242	21.258	19	0.89	2.67	0.34
27		65340	23.849	21	0.88	2.63	0.34
28		47597	17.373	15	0.86	2.74	0.31
ROAD SECTIONS							
29	1.5	26869	14.711	43	2.92	2.18	1.34
30	0.9	18337	6.024	27	4.48	3.35	1.34
31	1.5	29614	16.214	57	3.52	2.77	1.27
32	1.3	33143	15.726	45	2.86	2.78	1.03
33	1.6	35132	20.517	33	1.61	2.68	0.60
34	2.5	25577	23.339	29	1.24	2.63	0.47
35	3.5	34892	44.575	43	0.96	2.46	0.39
36	2.6	36254	34.405	24	0.70	2.52	0.28
37	2.0	53895	39.343	17	0.43	2.49	0.17

gory, which emphasizes the positive safety effect of traffic control signals.

Calculations in agreement with the accident seriousness method are presented in Table 2. The table exhibits the total number of accidents, deaths, and injuries that resulted at each site during the study period. The ETAN values presented in the table were calculated using the following equation

$$\text{ETAN} = aF + bJ + \text{TAN} \tag{2}$$

where

F = number of persons who died on the site,
J = number of persons injured on the site,
TAN = total accident number on the site, and
a and b = calibration factors.

The factors a and b were calibrated by trial and error. Values from 9 to 15 were assigned for a, and values from 3 to 5 were assigned for b. The ETAN was calculated for all possible combinations of a and b values. After each operation, the sites were ranked in ascending order according to their ETAN values. However, no significant differences were found between the ranks for the values of a and b within the range presented. The best values were obtained from the operation that gives closer ranks to those based on the other three methods of rating. It was found that $a = 12$ and $b = 3$ are typical values. The software package Lotus 1–2–3 was used to facilitate the computations.

Table 3 presents the ranks of the sites according to the four accident rating methods. The four ranks of each site were added to produce the danger index (DI). Priority or final ranking was then obtained on the basis of the DI. There is no significant difference between the ranks $R3$ and $R4$ because the danger factor (on which $R4$ is based) is dependent on the accident rate (on which $R3$ is based). The table also indicates that the final ranking, which is based on the DI, agrees better with $R3$ and $R4$. As mentioned previously, this finding agrees with the discussion on Table 1, which indicated that sites of higher accident frequency are not more dangerous. This evidence is focused by examining the second and last columns in Table 3. Clear examples are Sites 12, 20, and 27.

TABLE 2 ACCIDENT SERIOUSNESS AND ETAN VALUES

SITE NO.	LENGTH (KM)	ACC.NO. (TAN)	FATALITIES (F)	INJURIES (J)	ETAN	RANK
			GROUP 1			
6		50	0	26	128	1
9		22	3	22	124	2
2		33	0	20	93	3
1		33	1	16	93	4
7		18	0	18	72	5
4		34	0	11	67	6
12		25	0	9	52	8
8		20	2	4	56	7
3		24	0	7	45	9
5		20	0	6	38	10
11		12	0	8	36	11
10		16	0	5	31	12
16		12	0	5	27	13
14		20	0	2	26	14
13		13	0	3	22	15
15		10	0	2	16	16
17		10	0	2	16	17
			GROUP 2			
18		29	3	39	182	1
19		25	0	13	64	2
22		27	0	10	57	3
27		21	0	7	42	4
24		23	0	4	35	5
26		19	0	5	34	6
21		16	0	5	31	7
23		17	0	3	26	8
25		20	0	2	26	9
20		16	0	2	22	10
28		15	0	2	21	11
			ROAD SECTIONS			
32	1.3	45	1	24	99	1
30	0.9	27	2	10	90	2
31	1.5	57	1	21	88	3
29	1.5	43	2	19	83	4
33	1.6	33	2	11	56	5
35	3.5	43	1	18	31	7
37	2.0	17	1	11	31	8
34	2.5	29	2	8	31	6
36	2.6	24	0	7	17	9

ETAN=12F+3J+TAN For groups 1&2,
For road sections divide by section length in KM .

The DI is used to identify and rank the sites according to their danger significance. It is a measure for a site within the same group and should not be used to compare sites in two different groups. The DI is not a scale of seriousness of the sites within the group; for example, a site having DI = 50 is not five times more dangerous than a site having DI = 10. The only purpose for determining the DI is to use this index for final ranking of the sites for future developments. Thus, the sites within the groups were ranked in ascending order according to their DI. The site that had the lower value of DI was considered more dangerous and was given priority for funding.

CONCLUSIONS

Identification of accident-prone locations is a significant issue that should not be ignored in any traffic safety program. Such locations have been identified in Greater Amman. Sites for which improvements are imperative were also pointed out. Traffic volume studies revealed some significant conclusions that are useful for future studies. However, the following conclusions can be drawn from this investigation:

- Identification of accident-prone locations by the accident frequency method alone may lead to improper judgment.
- Accident-prone locations can be best identified by the quality control method (critical rate).
- Combining more than one method of accident rating is an appropriate technique for priority ranking.
- Seasonal expansion factors indicate that April traffic volume represents the average monthly traffic volume of the year.

ACKNOWLEDGMENTS

This work is part of a traffic safety project for Jordan sponsored by Mu'tah University. The author is grateful for the sponsorship and research facilities offered by the university. The author is also indebted to the Public Security Directorate, Traffic Police Department, for providing the accident information.

TABLE 3 RANKING OF SITES ACCORDING TO ACCIDENT RATES AND PRIORITY FOR IMPROVEMENT

SITE NO.	FREQUENCY R1 (1)	ETAN R2 (2)	RATE R3 (3)	DF R4 (4)	DANGER INDEX DI (1+2+3+4)	PRIORITY RANKING
GROUP1						
1	3	4	1	1	9	1
2	4	3	3	2	12	2
6	1	1	6	6	14	3
4	2	6	5	4	17	4
3	6	9	2	3	20	5
5	8	10	4	5	27	6
9	7	2	9	9	27	7
7	11	5	7	7	30	8
8	9	7	8	8	32	9
12	5	8	14	12	39	10
10	12	12	12	10	46	11
11	14	11	11	11	47	12
14	10	14	15	14	53	13
13	13	15	13	13	54	14
15	16	16	10	15	57	15
16	15	13	16	16	60	16
17	17	17	17	17	68	17
GROUP2						
18	1	1	1	1	4	1
19	3	2	2	2	9	2
22	2	3	5	5	15	3
24	4	5	7	7	23	4
20	10	10	3	3	26	5
21	9	7	4	4	24	6
23	8	8	6	6	28	7
25	6	9	8	8	31	8
26	7	6	9	9	31	9
27	5	4	11	11	31	10
28	10	11	10	10	41	11
ROAD SECTIONS						
29	3	4	1	1	9	1
31	1	3	3	3	10	2
32	2	1	4	4	11	3
30	7	2	2	2	13	4
33	5	5	5	5	20	5
34	6	6	6	6	24	6
35	4	7	7	7	25	6
36	8	9	8	8	33	8
37	9	8	9	9	35	9

REFERENCES

1. T. S. Khedaywi, A. H. Balbissi, and A. I. Abu-Seikh. Analysis of Automobile Accidents in Jordan. *ITE Journal*, Vol. 57, Dec. 1987, pp. 46–48.
2. *The Hashemite Kingdom of Jordan Statistical Yearbook*, No. 39. Department of Statistics, Amman, Jordan, 1988.
3. A. Gharaybeh. *Traffic Accidents in the Capital Governate*. Traffic Police Department Publications, Amman, Jordan, 1988.
4. I. Ismail. An Overview of Traffic Safety in Jordan. *ITE Journal*, Vol. 58, Nov. 1988, pp. 43–46.
5. *Annual Report*. Jordan Petroleum Refinery Company, Amman, Jordan, 1988.
6. M. J. Maher and L. J. Mountain. The Identification of Accident Blackspots: A Comparison of Current Methods. *Accident Analysis and Prevention*, Vol. 20, No. 2, 1988, pp. 143–151.
7. M. Plass and W. Berg. Evaluation of Opportunity-Based Rate Expressions. In *Transportation Research Record 1111*, TRB, National Research Council, Washington, D.C., 1987, pp. 42–48.
8. T. Chira-Chavala and K. K. Mak. Identification of Accident Factors on Highway Segments: A Method and Applications. In *Transportation Research Record 1068*, TRB, National Research Council, Washington, D.C., 1986, pp. 52–58.
9. J. L. Higle and J. M. Witkowski. Bayesian Identification of Hazardous Locations. In *Transportation Research Record 1185*, TRB, National Research Council, Washington, D.C., 1988, pp. 24–36.
10. M. Al-Esa, G. Ergun, S. Al-Senan, and A.-R. Al Zahrani. Development of a Hazardousness Index for Urban Intersections. *Proc., 3rd IRF Middle East Regional Meeting*, Riyadh, Saudi Arabia, Vol. 2, 1988.
11. D.-Y. Lin. Establishment of Road Accident Management Information System and Performance Evaluation of Accident-Prone Locations Improvement Programs in Taiwan, Republic of China. *Proc., 3rd IRF Middle East Regional Meeting*, Riyadh, Saudi Arabia, Vol. 2, 1988.

Publication of this paper sponsored by Committee on Traffic Records and Accident Analysis.

Highway Accidents: A Spatial and Temporal Analysis

WILLIAM R. BLACK

A temporal, spatial, and spatial-temporal autocorrelation analysis of highway accidents on the Indiana Toll Road from 1983 to 1987 is presented. Applications of von Neumann's ratio, Moran's I, nearest-neighbor analysis, and a spatial-temporal autocorrelation coefficient to a transportation network situation are illustrated. Applications of these methods to transport network attributes, such as accidents, have not appeared previously. The main objectives are to determine whether these techniques are sensitive enough to distinguish different patterns in the accident distributions and whether these patterns are explainable. The analysis involved 10 sets of accident data, categorized by date of occurrence and location on an east-west roadway. Only 2 of the 10 revealed positive temporal autocorrelation (clustering in time), 5 revealed positive spatial autocorrelation (clustering in space), and between 6 and 9, depending on the method used, revealed positive spatial-temporal autocorrelation (clustering in time and space). These results suggest that observed autocorrelations in accidents are a function of weather conditions or traffic volumes, or a combination of the two.

The spatial, temporal, and spatial-temporal distribution of motor vehicle accidents along a major highway is examined. Although temporal analyses have been undertaken for several years as analysts have sought to predict the number of accidents or fatalities on highways, methods for rigorously analyzing the spatial distribution of events along a highway are not common. Methods for analyzing the spatial and temporal distributions simultaneously on a linear network have not been presented before. Two approaches to this latter problem are examined here. One of these is drawn from the plant ecology and geographical analysis literature and the other has its origin in epidemiology.

CLUSTERING IN TIME AND SPACE

Motor vehicle accidents should occur as a random series in time or space. However, these incidents often cluster temporally and spatially. The existence of temporal clustering is evident during holiday weekends, when the number of accidents and fatalities increases in response to heavier-than-usual traffic volumes. Whether spatial clustering exists is not as obvious.

Temporal Clustering

To determine whether clustering is occurring in a temporal sense, it is necessary to have some notion of an appropriate time interval for data analysis. If temporal distribution is examined at intervals of less than 1 min, the only events to cluster might be those reflecting the same vehicle-to-vehicle collisions and other single-vehicle random events. This information is not without interest, but it does not suggest much to the analyst or policy maker. Therefore, a different temporal scale of analysis is necessary.

The time interval used here is the day. Its use is based on the belief that the number of highway accidents today is, in part, related to the number of accidents that occurred yesterday. Cause is not implicit in such a statement, but rather the recognition that daily accidents are often related to traffic flow volumes or weather conditions, which vary daily, weekly, and seasonally. As a result, temporal autocorrelation exists in these data with variable levels for 1 day not differing a great deal from the level of that variable the following day; that is, similar values tend to cluster in time. Use of a day for the time interval should pick up the influence both of flow volumes and of weather conditions.

Of course, if the major factors influencing temporal autocorrelation are absent, there should not be any clustering of the events in time. For flow volumes and weather conditions, it is possible that increased patrolling by state police and increased maintenance (such as snow removal) could offset some of the expected temporal clustering. Nevertheless, a tendency toward temporal clustering is expected.

Spatial Clustering

Spatial autocorrelation is the tendency for the level of a variable at one location to influence the level of that variable at sites in proximity to the first location. If positive spatial autocorrelation is present, it results in a spatial clustering of similar variable values. This clustering may be caused by many factors. Among these factors are higher traffic volumes in different regions of the transport line, natural or anthropogenic environmental factors (such as fogs) that restrict or interfere with vehicle operation, points of access or egress at which vehicle speeds change, or areas of poor highway design in an engineering sense. A clustering of similar accident values in space would be expected if one or more of these factors are present.

The proper spatial interval for the analysis of spatial autocorrelation in this context is debatable. As explained previously, setting a small interval will result in the clustering only of vehicles in the same incident (automobile-to-automobile collisions), but these data are not of interest here. The use of 0.5-km or similar intervals would be desirable for

Department of Geography, Indiana University, Kirkwood Hall 202, Bloomington, Ind. 46405.

certain environmental events, but this level of detail is uncommon and, as a result, not available. Because of data reporting standards, this analysis uses the 1-mi segment. Most major regional phenomena should be perceptible at this scale, and minor clustering at a lower scale should also be identifiable at the larger scale.

Spatial-Temporal Clustering

Because a scale is available for locating every incident in time and another scale is available for measuring every incident in space, the question is whether there is some method of combining these scales to determine if certain incidents are clustering in time and space. If such a method were available, it might enable researchers to evaluate whether the temporal clustering and spatial clustering were due to the same or different major events. For example, assume that the temporal clustering found in a series of automobile accidents was attributable to holiday travel on weekends during the winter. The presence of spatial clustering might suggest a concentration of this same series of accidents in a given region. The presence of spatial-temporal clustering would reveal that accidents were occurring on the same weekends and sections of the highway and that the two separate distributions are actually a single, interrelated, spatial-temporal distribution. The objective is to identify a method for analyzing this latter type of distribution.

If 1-day time intervals are acceptable for the measurement of temporal autocorrelation and 1-mi spatial intervals are equally acceptable for measuring spatial autocorrelation, then it is reasonable to use these two dimensions to define an area called time-space. In time-space, all events are identifiable as occurring within a certain time period and within a bounded space. If the initial metrics for time and space are unreasonable, the analysis of a time-space with dimensions defined by these metrics will be of little value.

METHODOLOGY

Temporal Autocorrelation Analysis

Statistics for assessing temporal autocorrelation in a data series include the Durbin-Watson statistic and the von Neumann ratio (*1*, pp. 305–307). The von Neumann ratio (Q) was chosen because it was easier to evaluate. For a set of n observations on some variable x arranged in a successive series, the statistic is calculated as follows:

$$Q = \frac{[1/(n-1)] \sum_{i=2}^{n} (x_i - x_{i-1})^2}{(1/n) \sum_{i=1}^{n} (x_i - \bar{x})^2} \tag{1}$$

For a large number of observations ($n > 60$), the expected value of Q follows a normal distribution with mean $2n/(n-1)$ and variance $4/n$. Evaluation of the ratio is accomplished in the traditional manner using standard normal deviates on the basis of these values and the calculated Q.

Spatial Autocorrelation Analysis

Analysis of spatial autocorrelation for a linear system, such as a highway, involves the adaptation of conventional autocorrelation techniques used in the analysis of point and areal distributions to linear situations. The statistic of choice in such cases is most often Moran's I (*2–4*). The statistic is calculated as follows:

$$I = \frac{n}{\sum\sum w_{ij}} \frac{\sum\sum w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum(x_i - \bar{x})^2} \tag{2}$$

where

x_i = value of variable x on Line Segment i,
$\bar{x}$ = mean of variable x,
n = number of line segments, and
w_{ij} = weights indicating whether Line Segment i is connected to Line Segment j (= 1) or not (= 0).

For continuous data, the expected value of Moran's I is

$$E(I) = \frac{-1}{n-1} \tag{3}$$

The variance of I under the assumption of normally distributed data is

$$\mathrm{Var}(I) = \frac{n^2 S_1 - n S_2 + 3\left(\sum\sum w_{ij}\right)^2}{\left(\sum\sum w_{ij}\right)^2 (n^2 - 1)} - E(I)^2 \tag{4}$$

where

$$S_1 = \frac{1}{2} \sum\sum (w_{ij} + w_{ji})^2 \tag{5}$$

and

$$S_2 = \sum\left(\sum w_{ij} + \sum w_{ji}\right)^2 \tag{6}$$

Although accident data of the type used here are not usually normally distributed, the presence of a large number of observations in the series analyzed makes this assumption tenable. The difference between the expected value of I (from Equation 3) and the calculated value of I (from Equation 2) may be divided by the square root of the variance (see Equation 4) to yield a standard normal deviate, often called a z score, for hypothesis testing.

Spatial-Temporal Autocorrelation Analysis

Given that it is possible to assess the autocorrelation of the accident data in time and one-dimensional space, the influence of these factors is examined simultaneously. There are two ways that the data could vary. First, the pattern of cases in an area defined by the space and time dimensions can be examined. This method is known as pattern analysis, and the finding of a clustered pattern would infer some type of dependence in the distribution. Second, the extent to which the pattern created by values (in this case, the number of accidents) displays some evidence of organization can be studied. This method is spatial autocorrelation analysis, and a finding of positive spatial autocorrelation would also suggest some type of dependence in the distribution.

A common method for measuring whether clustering exists in a two-dimensional spatial pattern of cases is known as nearest-neighbor analysis. Such a technique could also be used to examine the pattern of cases in time-space. Nearest-neighbor analysis was developed primarily by the plant ecologists Clark and Evans (*5*), although geographers have done a considerable amount of work on this subject since that early research (*6–10*). The measure compares the actual distance of each point to its nearest neighbor in two-dimensional space, with the expected distance to the nearest neighbor being based on a random distribution of points in that space. Division of the former by the latter yields an index for which unity indicates a random distribution, zero represents complete clustering, and a completely dispersed pattern yields a value of approximately 2.15.

Hypothesis testing proceeds by finding the difference between the observed and expected nearest neighbor distances and dividing this value by the expected standard error for a random distribution to yield an index, which is also a standard normal deviate. Symbolically, the observed average distance to the nearest neighbor ($\bar{d}$) is

$$\bar{d} = \sum d_{ij\min}/n \tag{7}$$

where $d_{ij\min}$ is the distance of i to its nearest jth neighbor, and n is the number of points in the distribution. The expected average distance for a random distribution ($\bar{d}_{\text{ran}}$) is

$$\bar{d}_{\text{ran}} = 1/[2(n/A)^{0.5}] \tag{8}$$

where A is the area of surface occupied by the point distribution. The standard error of the mean nearest-neighbor distances is

$$\text{SE}_{\bar{d}} = 0.26136/[n(n/A)]^{0.5} \tag{9}$$

and the standard normal deviate in this case is

$$z = \frac{\bar{d} - \bar{d}_{\text{ran}}}{\text{SE}_{\bar{d}}} \tag{10}$$

Methods for evaluating the pattern of values in space and time fall within the domain of spatial autocorrelation analysis. The methods are not that well developed for time and two-dimensional space. However, for one-dimensional time and one-dimensional space (such as the toll road of interest here), it is possible to use a cross-product statistic attributed to Knox (*11*) and Mantel (*12*).

The technique is described in some detail by Cliff and Ord (*8*) and Upton and Fingleton (*10*). It involves in this case the construction of two event matrices. If there are n events, then the matrices are $n \times n$. For the first matrix (the time matrix), a 1 is included in some Cell t_{ij} if Event i occurred within one time unit of Event j, and 0 otherwise. The second, or space, matrix includes a 1 in Cell s_{ij} if the events occurred within one unit of each other in space, and 0 otherwise. For both matrices, if $i = j$, then the entry is 0. The general cross-product statistic is obtained from the following equation:

$$R = \sum_{i=1}^{n} \sum_{j=1}^{n} t_{ij} s_{ij} \tag{11}$$

If the events are completely independent, then $R = 0$. The expected value of R is

$$\text{E}(R) = \sum_{i=1}^{n} \sum_{j=1}^{n} t_{ij} \sum_{i=1}^{n} \sum_{j=1}^{n} s_{ij}/n(n - 1) \tag{12}$$

Upton and Fingleton (*10*, p.157) provide details on the calculation of variance. The obtained R value can be evaluated with the test statistic, as follows:

$$z = \{[R - \text{E}(R)] - 1\}/[\text{var}(R)]^{0.5} \tag{13}$$

ILLUSTRATIONS

Table 1 presents an illustration of these methods. The first three variables represent three different distributions of 17 events across 40 time periods. The concern goes to the nature of these three distributions. As presented in Table 2, use of the von Neumann ratio reveals that the first variable represents significant positive temporal autocorrelation, the second represents significant negative temporal autocorrelation or dispersion, and the third does not display a significant level of temporal autocorrelation.

The second three variables of Table 1 represent the locations of 17 events along a single dimension, such as a highway, with a length of 40 units. Spatial autocorrelation analyses of these variables also suggest three different types of distributions (see Table 2). The first has a high level of positive spatial autocorrelation (i.e., a clustering of similar values), the second has a high level of dispersion or a uniformity in the distribution of events, and the third is neither clustered nor dispersed but tends toward a random distribution.

Figure 1 shows several possible mappings of the temporal and spatial variables in time-space. In each case the mappings are only two of the possible arrangements that could result when an event's temporal location is linked with its corresponding spatial location. By referring to the marks on the horizontal and vertical axes of the figures, the nature of the one-dimensional temporal and spatial autocorrelation (analogous to the pattern) is revealed. The figure shows that the presence of positive spatial autocorrelation in the location variable and positive autocorrelation in the temporal variable may lead to a clustered distribution (1a) or a dispersed distribution (1b). Furthermore, a dispersed or uniform distribution for these two axes may also lead to a clustered pattern (2a) or a dispersed pattern (2b). Finally, a nearly random distribution of these two axes may result in a clustered pattern (3a) or a dispersed pattern (3b) of incidents in time-space.

Table 3 presents the results of a nearest-neighbor analysis of each of these six mappings in time-space. The nearest-

TABLE 1 SIX EXAMPLES OF SPATIAL AND TEMPORAL LOCATION OF EVENTS

Day or	Temporal			Spatial		
Mile	1	2	3	1	2	3
1	0	0	0	0	0	0
2	0	0	1	1	1	1
3	0	0	0	1	0	0
4	0	1	0	1	1	0
5	0	0	0	1	0	0
6	1	1	0	0	1	0
7	1	0	1	0	0	1
8	1	1	0	0	1	0
9	1	0	0	0	0	0
10	0	1	1	0	1	1
11	0	0	1	0	0	1
12	0	1	1	0	1	1
13	0	0	0	0	0	0
14	1	1	1	1	1	1
15	1	0	0	1	0	0
16	1	1	0	1	1	0
17	1	0	1	1	0	1
18	0	1	1	0	1	1
19	0	0	0	0	0	0
20	1	1	0	0	1	0
21	1	0	0	0	0	0
22	1	1	1	0	1	1
23	1	0	0	0	0	0
24	0	1	0	1	1	0
25	0	0	1	1	0	1
26	0	1	0	1	1	0
27	0	0	1	1	0	1
28	0	1	1	1	1	1
29	0	0	0	0	0	0
30	0	1	0	0	1	0
31	0	0	1	0	0	1
32	0	1	0	0	1	0
33	0	0	0	0	0	0
34	0	1	1	1	1	1
35	0	0	0	1	0	0
36	1	1	0	1	1	0
37	1	0	0	1	0	0
38	1	0	1	0	0	1
39	1	0	1	0	0	1
40	1	0	1	0	0	1

neighbor analysis confirms what is visually apparent from Figure 1. Distributions 1a, 2a, and 3a are clustered, and distributions 1b, are dispersed in the time-space. In other words, the patterns of Mappings 1a, 2a, and 3a reveal the presence of positive spatial autocorrelation, and the patterns of Mappings 1b, 2b, and 3b indicate its absence.

Table 3 also presents the results of a Knox R analysis of these six situations. In so doing, it identifies what might be the major shortcoming of this technique. Although the Knox R indicates that Distributions 1a and 3a have positive spatial autocorrelation, the test completely fails for Distribution 2a. Although clustering is apparent in this distribution, each case

TABLE 2 ILLUSTRATIVE TEMPORAL AND SPATIAL CASES

Temporal Data Sets	von Neumann Q	z
Distribution 1	.7345	-4.16*
Distribution 2	3.5674	4.79*
Distribution 3	2.2034	.48
Spatial Data Sets	**Moran's I**	**z**
Distribution 1	.5870	3.92*
Distribution 2	-.8601	-5.35*
Distribution 3	-.1029	- .50

* significant at .01 ($z > |2.58|$)

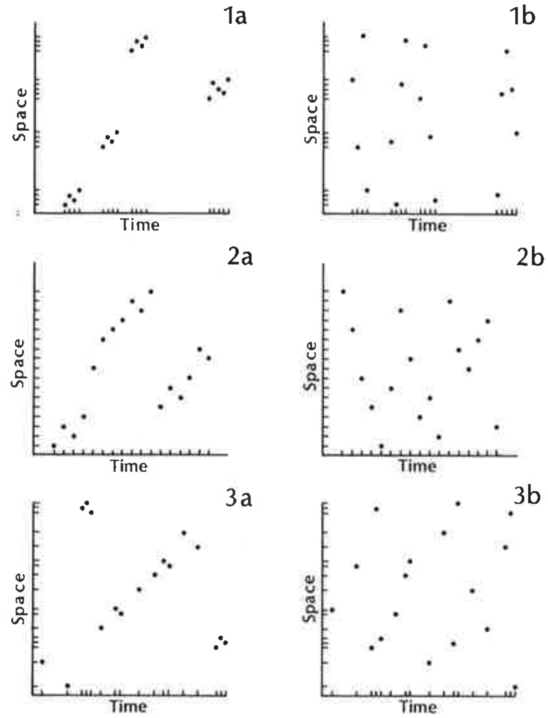


FIGURE 1 Possible two-dimensional mappings of spatial and temporal events.

is more than one unit away from each other case in time and space. This finding demonstrates how critical the definition of nearness is in using this test. The other distributions (1b, 2b, and 3b) reveal an absence of positive spatial-temporal autocorrelation, as expected.

Of the various distributions, the ones that are of primary interest for analysis or policy reasons are those that display positive spatial autocorrelation in one dimension, positive temporal autocorrelation, and positive spatial-temporal autocorrelation, or a clustering in time-space, as assessed by nearest neighbor analysis. Random distributions would imply no major regional or temporal influences. A dispersed or nearly uniform pattern is an unlikely, though not impossible, distribution for most transport-related phenomena.

STUDY AREA AND DATA

The empirical situation of interest here is the spatial and temporal distribution of motor vehicle accidents along the Indiana Toll Road from 1983 through 1987. This toll road is a limited-access trafficway in the northern part of Indiana extending 156 mi between Illinois and Ohio (see Figure 2). The trafficway consists of an eastbound roadway and a westbound roadway.

The motor vehicle accident data are perhaps best thought of as traffic incidents. The incidents recorded may be minor and involve suspected personal injury or damage to a single item (e.g., a vehicle, a sign, or a guardrail) in excess of $500. The data also record incidents involving fatalities, but these reports are a small minority of those in this data base. The data used are recorded by day of the year and toll road milepost. The former allows an analysis of temporal autocorrelation over 365 days (366 days in 1984), whereas the latter permits a spatial autocorrelation analysis across the 156 days.

For the spatial-temporal analysis, the two scales above defined a two-dimensional time-space surface. On this surface an accident receives a space coordinate and a time coordinate corresponding to the milepost and day on which the accident occurred. Events that occurred near each other in time and space will appear close to each other on the time-space surface. If the analysis suggests a clustering of events, it is reasonable to infer the existence of a certain amount of spatial-temporal autocorrelation.

RESULTS OF ANALYSIS

The results of the temporal autocorrelation analysis are presented in Table 4 as von Neumann Q statistics and standard normal deviates. Of the 10 data sets analyzed, only 2 demonstrated a significant amount of temporal autocorrelation. The implications of this finding are numerous and range from very successful patrolling practices to uniqueness in the type of data used. It is common practice for the Indiana State Police to increase the level of patrolling during holiday periods in an attempt to decrease motor vehicle accidents. The toll road is not a typical highway in that it receives heavier use during

TABLE 3 ILLUSTRATIVE SPATIAL-TEMPORAL CASES

Time-Space Data Sets	Nearest Neighbor Index	z	Knox Statistic	z
Mapping 1a	.386	- 4.85*	10	3.09*
Mapping 1b	1.374	2.95*	0	.71
Mapping 2a	.725	- 2.17**	0	0.00
Mapping 2b	1.375	2.96*	0	0.00
Mapping 3a	.602	- 3.14*	12	10.56*
Mapping 3b	1.466	3.68*	0	- .47

* significant at .01 (z > 2.58)

** significant at .05 (z > 1.96)

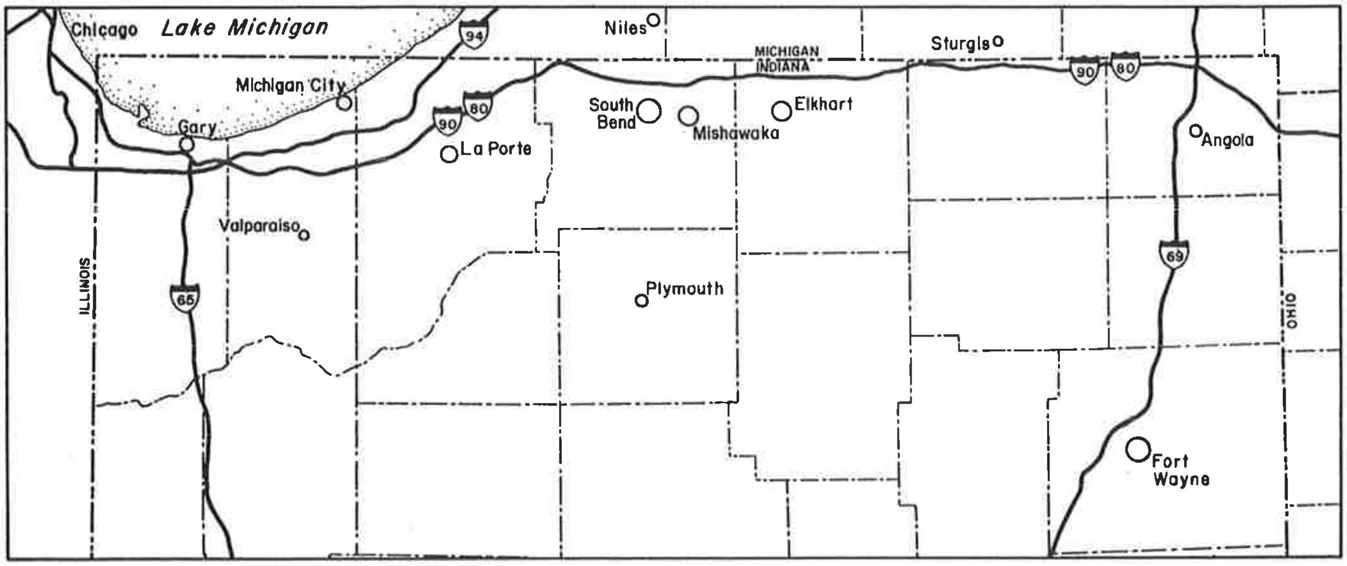


FIGURE 2 Location of Indiana Toll Road (I-80 and I-90) in northern Indiana.

weekdays with commuters and truckers, and relatively light weekend traffic. Nevertheless, a temporal clustering during weekdays is also not apparent.

The spatial autocorrelation analysis did reveal a significant amount of positive spatial autocorrelation (at the .05 significance level) for 5 of the 10 data sets analyzed. This result may reflect an underlying influence of heavier traffic volumes in the western part of the state due to the Chicago metropolitan area. It may also represent a greater number of accidents due to the mixing of traffic moving at different speeds, because the western part of this highway has a greater number of entrances and exits from the toll road than does the eastern end.

The spatial-temporal analysis consisted of a nearest-neighbor analysis of the motor vehicle accidents on a time-space surface and a spatial temporal autocorrelation analysis using Knox's statistic. The results of the nearest-neighbor analysis are presented in Table 5, which reveals a significant amount of clustering in 6 of the 10 data sets analyzed. Assuming comparability of the z scores between the one-dimensional spatial and temporal autocorrelation and the two-dimensional nearest-neighbor analysis, six of the two-dimensional cases

TABLE 4 VON NEUMANN'S Q STATISTIC, MORAN'S I STATISTIC, AND CORRESPONDING z SCORES

Year	Indices	Eastbound	Westbound
1983	Q (z)	1.80 (-2.01)**	1.84 (-1.57)
	I (z)	.1270 (1.67)	.1573 (2.05)**
1984	Q (z)	1.76 (-2.33)**	1.87 (-1.29)
	I (z)	.2098 (2.71)*	.1219 (1.60)
1985	Q (z)	1.95 (- .49)	1.85 (-1.47)
	I (z)	.1638 (3.13)**	.3177 (3.81)*
1986	Q (z)	1.84 (-1.62)	2.01 (.01)
	I (z)	.0862 (1.16)	.1133 (1.50)
1987	Q (z)	2.02 (.17)	1.90 (-1.02)
	I (z)	.1505 (1.97)**	.0789 (1.07)

* significant at .01 (z > 2.58)

** significant at .05 (z > 1.96)

TABLE 5 SPATIAL-TEMPORAL NEAREST-NEIGHBOR AND KNOX *R* ANALYSIS

Year	Roadway	n	NN Index	z	Knox-R	z
1983	East	337	.915	-2.99*	66	5.11*
1983	West	290	.877	-4.00*	46	3.57*
1984	East	323	.884	-4.00*	64	4.62*
1984	West	311	.964	-1.22	56	4.91*
1985	East	368	.938	-2.26**	68	4.15*
1985	West	353	.891	-3.92*	76	4.98*
1986	East	335	1.007	.26	24	.03
1986	West	340	.910	-3.17*	64	5.55*
1987	East	394	.973	-1.03	52	2.77*
1987	West	335	.976	- .83	52	3.71*

* A significant z for a two-tailed test at the .01 level is > ¦2.58¦.

** A significant z for a two-tailed test at the .05 level is > ¦1.96¦.

are stronger than their one-dimensional counterparts. This finding is to some extent consistent with the combining of the low-level temporal autocorrelation with the low one-dimensional positive spatial autocorrelation.

The level of temporal autocorrelation was the highest for roadways East-83, West-83, East-84, West-85, and East-86, although only two of these were significant. Of these five, the first four are strong in the two-dimensional analysis. Eastbound traffic in 1985 was having incidents that significantly clustered in space, but not time, and combining the two increased the level of clustering. Westbound traffic in 1986 that was not significantly clustered in time or space was significantly clustered in time-space. In four cases (West-84, East-86, East-87, and West-87) the nearest-neighbor analysis revealed no significant clustering in time-space.

Plots of three of these distributions are shown in Figures 3–5, where each box represents a traffic incident. The eastbound traffic incidents of 1983 are plotted in Figure 3. Notable attributes of this figure are (a) the vertical array of symbols at approximately Day 80 across the entire length of the tollway, (b) the clustering of incidents between Mileposts 0 and 25 through the entire year, (c) the vertical array of incidents between Days 350 and 365, and (d) a clustering of incidents around Day 150 between Mileposts 90 and 100. Of these four attributes, the first and third appear to be a function of snow and ice conditions according to climatological records, and the second appears to be related to the generally higher traffic volume near Chicago (which increases the probability of an incident). The reason for the clustering noted at Day 150 is not apparent, although its occurrence on Memorial Day suggests a traffic volume situation. The first three patterns are also evident in the westbound traffic of 1983 (see Figure 4), which is consistent with the regional weather pattern explanation offered. The fourth pattern does not appear in the westbound traffic.

As noted previously, regional weather conditions play a major role in the development of spatial clustering. Snowfall, snow coverage, ice, and rainfall are evident in the mapped data. There was also a period of several days in October 1987 when a clustering of incidents occurred near South Bend.

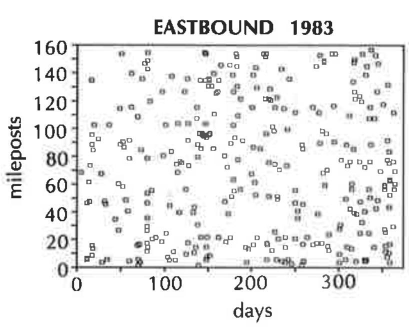


FIGURE 3 Time-space plot of accidents on Indiana Toll Road: eastbound roadway, 1983.

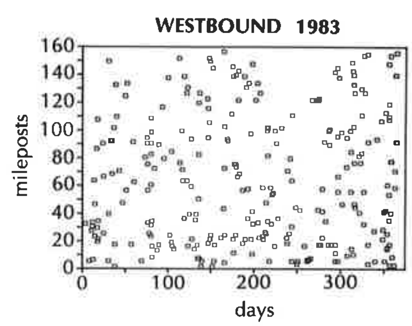


FIGURE 4 Time-space plot of accidents on Indiana Toll Road: westbound roadway, 1983.

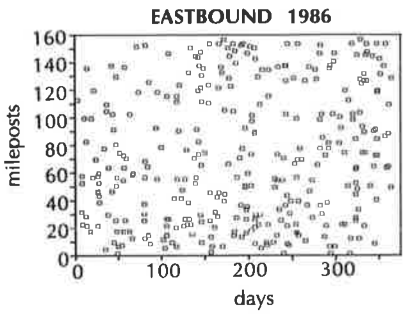


FIGURE 5 Time-space plot of accidents on Indiana Toll Road: eastbound roadway, 1986.

Local climate records indicate a record number of fog days during that month.

Figure 5 shows the eastbound traffic of 1986. It is included primarily because it represents a random distribution in time-space. Large areas of this time-space have no incidents, whereas in other areas there is an occasional overlapping of two or, on rare occasions, three incidents. Such a pattern is typical of a random two-dimensional distribution. From a safety analyst's perspective, this pattern is a desirable one.

Table 5 also presents the results of the Knox R spatial-temporal analysis of accident values. These results differ substantially from those of the nearest neighbor analysis and suggest a much stronger level of spatial-temporal autocorrelation in the pattern of values. The two methods are in perfect agreement in the nearly random pattern of data for eastbound traffic of 1986, but otherwise the Knox R analysis indicated highly significant spatial-temporal autocorrelation in every case.

The spatial-temporal results should not be viewed as conflicting. The nearest-neighbor analysis is examining the pattern of events in time-space. The Knox analysis is examining the similarity (autocorrelation) in the pattern of events in time and the pattern of events in space. From a residual analysis perspective, the Knox approach may be more useful; from an accident analysis perspective, the nearest-neighbor approach seems to yield more useful results.

POLICY IMPLICATIONS

Each of the situations analyzed implies the need for or the success of some public safety policy. Whether or not the traffic on the Indiana Toll Road is typical, the absence of significant temporal autocorrelation in motor vehicle accidents is a desirable attribute. This success may be attributable to policing, good road maintenance (snow removal or highway surface maintenance), or a good traffic use pattern.

The presence of spatial clustering implies that all is not perfect with this system. Weather conditions in given regions may create the spatial clustering of accidents, but clustering would then be expected to appear in the temporal analysis and it does not. The clustering may simply result from increases in accidents because of increases in traffic volume. One possible policy response might be to use on- and off-ramps to control traffic volumes. The spatial clustering may also be caused by poor highway design.

The time-space analyses, as reflected by the nearest-neighbor analyses, suggest only a few instances in which a tendency toward temporal clustering is combining with spatial clustering to make insignificant concentrations appear significant. The implications might be that poor weather conditions over a few time periods are combining with poorly designed ramps and resulting in the identification of a stronger level of clustering. The ability to identify such situations suggests that this approach to time-space analysis deserves further study.

ACKNOWLEDGMENTS

The author wishes to express his appreciation to John M. Hollingsworth for preparing the figures and to Dawn E. Hewitt for assisting in the production of Figures 3–5.

REFERENCES

1. H. Theil. *Introduction to Econometrics*. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1978.
2. M. F. Goodchild. Spatial Autocorrelation. *Concepts and Techniques in Modern Geography*, No. 48, 1987.
3. D. A. Griffith. *Spatial Autocorrelation: A Primer*. Association of American Geographers, Resource Publications in Geography, Washington, D.C., 1987.
4. J. Odland. *Spatial Autocorrelation*. Scientific Geography Series, Vol. 9, Sage Publications, Newbury Park, Calif., 1988.
5. P. J. Clark and F. C. Evans. Distance to Nearest Neighbor as a Measure of Spatial Relationships in Population. *Ecology*, Vol. 35, 1954, pp. 445–452.
6. L. J. King. A Quantitative Expression of the Pattern of Urban Settlements in Selected Areas of the United States. *Tijdschrift Voor Economische en Sociale Geografie*, Vol. 53, 1962, pp. 1–7.
7. M. F. Dacey. Modified Poisson Probability Law for Point Pattern More Regular than Random. *Annals of the Association of American Geographers*, Vol. 54, 1964, pp. 559–565.
8. A. Cliff and J. Ord. *Spatial Processes*. Pion, London, 1981.
9. A. Getis and B. Boots. *Models of Spatial Processes*. Cambridge University Press, England, 1978.
10. G. Upton and B. Fingleton. *Spatial Data Analysis by Example: Point Pattern and Quantitative Data, Volume I*. Wiley, New York, 1985.
11. G. Knox. The Detection of Space Time Interactions. *Applied Statistics*, Vol. 13, 1964, pp. 25–29.
12. N. Mantel. The Detection of Disease Clustering and a Generalized Regression Approach. *Cancer Research*, Vol. 27, 1967, pp. 209–220.

Publication of this paper sponsored by Committee on Traffic Records and Accident Analysis.

Implementing a Distributed Responsibility Approach to Improving Data Quality

DAVID B. BROWN AND CAROLYN L. MCCREARY

An approach to improving data quality that is based on the concept of distributed responsibility is presented. This concept requires a direct distribution of data benefits throughout the data-responsible organizations. Premises on data quality are presented to formulate the problem so that the concept of distributed responsibility can be understood. Several accident data systems are discussed as examples of distributed systems, but proposals for continued development of this concept are strongly encouraged.

Data gathered on accidents, traffic citations, roadway characteristics, emergency medical services, and a host of other topics will be used in decision making whether or not the quality of the data is adequate. Although this statement certainly provides ample reason to improve data quality, a further analysis is required to obtain a full appreciation for the critical role played by data and the information derived therefrom in the decision-making process. Decision making in this sense refers to that process by which resources are allocated. The issues of traffic safety budget allocation are emphasized here, but the principles apply equally to any process in which scarce resources must be allocated to address a virtually infinite number of potential problems.

The statement that data will be used regardless of its quality is based purely on empirical observation of past practice. In the absence of other evidence on which to base a decision, the available data, albeit of the poorest quality, are generally given credence. Many pilot data collection projects suddenly turn into the real thing, and the data are used to prove hypotheses without any experimental design. Is this practice necessarily bad? Is it not possible that relatively good decisions have been made even though the data were not of the highest quality and were not gathered for the purposes to which they were later applied? These questions are addressed in the following paragraphs.

The only value of a vehicle crash is the information that it provides to decision makers to prevent a similar occurrence in the future. A crash costs anywhere from a few hundred to hundreds of thousands of dollars, depending on the severity. However, if it can be used to prevent other crashes, then it has intrinsic worth. However, this value may be completely lost if the quality of the data collected is inadequate. In fact, if the data are misleading, their use could be counterproductive. (This situation is not believed to be generally true, but practitioners should be aware of the possibility.) Finally, even good data that could be beneficial in guiding decisions can be neutralized by the questioning of its credibility. A series of premises are presented in the following section to set the stage for the proposed solution.

Department of Computer Science and Engineering, Auburn University, 107 Dunstan Hall, Auburn, Ala. 36849-5347.

PREMISES

Perfect Accuracy Is Impossible

The problems of data quality must be viewed in the proper perspective before attempting to formulate a conclusion. The major aspect of data quality is accuracy, defined in the context of traffic safety to include both completeness and validity of the data. Most traffic accident data elements are nominally coded data, that is, multiple choice questions. Invalid responses result from arbitrarily marking an incorrect response, misinterpretation of the question, data entry error, and other shortcomings of the data collection instrument itself. Incompleteness occurs when a data element is omitted, for whatever reason, and does not get recorded on the accident record. A second form of incompleteness occurs when a data element necessary for decision making does not exist in the data base. Another example of inaccuracy involves recording a milepost that varies from the actual accident location by a significant amount.

In a typical state data base of more than 100,000 traffic accident records per year, some inaccuracies are bound to creep in. The questions that should be asked are, What are the ramifications of these shortcomings? Does this inaccuracy alter policy to any degree? If not, there is no reason to worry about such errors. But if the deficiencies are altering policy—or, worse yet, if they are being used to discard the information value of the data altogether—then aggressive action is warranted to rectify these problems.

Perfect accuracy is unattainable, and those who insist on it are being unreasonable. No decision in life is based on total certainty. Why should the data on which safety decision making is based be any different? For example, election predictions are made with great accuracy on the basis of a small proportion of the results. If a good forecast can be made with less than 10 percent of the vote, should decision makers hesitate to use an accident records data base that is known to be at least 90 percent complete in most data elements? Of course, it is important that missing cases be randomly distributed throughout the population. The data base can then be viewed as a large random sample (proportionately speaking). If one geographic area or type of accident has more than its share of missing cases, adjustments can be made. However, the credibility of the remaining data does not automatically need to be questioned.

Another good example of sampling theory is in industrial quality control, which is a well-developed and sophisticated discipline. Products that require destructive testing for validation obviously cannot be subjected to 100 percent sampling. Sound inferences about the quality of a production run can be made with just a small percentage of the lot as a sample. This theory can certainly be applied in those states that monitor accidents by type as required by NHTSA for problem identification. Even though all accidents are not reported, sufficiently large numbers are available to ensure that the sample is relatively the same from year to year. Thus, when a particular type of accident (e.g., child pedestrian accidents in urban areas) is suddenly significantly higher in frequency than it was over the past few years, this situation can alert decision makers to take action.

Improvement in data quality will have little effect on the types of decisions for which current sampling is adequate. For these types of application (and possibly these only), the current quality is sufficient, and the data should be used and trusted. However, accident data bases are being pushed to produce more information of various types, which leads to the critical question, What quality is required? The answer depends on the application. The following table presents several applications with an assessment of the degree of accuracy required and available for each. These assessments are based on an estimate of the average state; any given state will probably vary from this norm. Clearly, a range of accuracy is required, and all data do not have the same quality requirement. With this concept in mind, it is important that the next premise be understood before proceeding.

Application	*Accuracy Assessment*
Problem identification	Currently adequate, Sufficient for Most Types
Hazardous location ID	Needs improvement but data are certainly usable
Location investigation	More details and greater accuracy would improve policy
Evaluation	Inadequate level of detail for many countermeasures

Too Much Data or Quality Can Be Counterproductive

Before the advocates of data quality become alarmed at this premise, they should recognize that too much of anything is, by definition, counterproductive. Thus, the argument reduces to determining whether or not it is possible to have too much data or too much data quality. To define terms, "too much data" refers to the collection of too many data elements, which can result in data collection forms that are too large or in data bases with too many records or too many types of records. These situations often occur in the name of completeness or because of the possibility that a certain application may occur in the future. The definition of quality, on the other hand, includes both accuracy and completeness. Finally, it is not being argued that there is currently too much data or that the quality is currently too high. The premise states only that there is a possibility that these two conditions can exist.

As stated, the purpose of collecting data is to guide the decision-making process in the allocation of limited resources. This objective leads to a paradox for the data base designer because data base design, data systems design, data collection, entry, maintenance, and ultimately processing to produce information all consume resources that could go into direct countermeasures. Could data (actually metadata, or data on data) be gathered to determine how many resources are to be allocated to the data subsystem? If so, should not an attempt be made to gather meta-metadata, or multimetadata? That this process of data collection could be extended infinitely proves the premise about data quantity to be true. The creation of even the first level of metadata is so expensive that most decisions about the content and structure of data bases are made on the basis of experience rather than hard data.

The data quality issue is more difficult to prove because there would seem to be an upper limit on data quality, especially for nominally coded variables, which make up most of the accident records data base. How much more difficult is it for an officer in the field to check the right code as opposed to the wrong code? How much more does it cost to measure the milepost to the nearest 0.01 mi as opposed to the nearest 0.1 mi? But it must be admitted that, no matter how hard one strives for perfect quality, it is always possible to do better. For example, the milepost reference system could be replaced with an alternative that would locate accidents much more accurately and reliably. Similarly, modifications could be made to most of the nominal-level variables to improve data quality. Are these improvements justified? In most cases, even minor enhancements cannot be made expediently because they will disrupt the system. Thus, compromises in quality are essential, which proves that there is an optimal level of quality and that it is somewhere on this side of perfection.

The practical ramifications of the premise are much more important than its proof. Too much data and an overemphasis on quality of certain unimportant data elements have resulted in a neglect of other data elements that are essential to decision making. When a data element is proposed, it should be recognized that the element must be properly defined, collected, entered, maintained, and processed if it is to be useful. A breakdown anywhere in this information-production chain can yield a data element that is counterproductive. Thus, an economic balance should be maintained, not just between data collection and direct countermeasures but also between competing data elements within a given data collection instrument.

The recognition that these two conditions (i.e., too much data and too much data quality) are attainable is important so that the objectives of the data-base design project can be put into perspective. Designers of data collection forms must recognize that they are responsible for the following: (a) determining the optimal number and type of data elements to maximize collector productivity, (b) designing these optimal data elements to produce the information required for effective decision making, and (c) providing guidance to the data collectors on the degree of accuracy required.

Although the proof of this premise would seem to oppose data quality advocates, such is not the intent. Rather, it is only by recognizing that an optimal level of quality exists that designers can deal with the necessary issues to achieve this level. The following premise addresses the current state of accident data as opposed to potential excesses.

Current Accuracy of Data Is Generally Not Sufficient

The qualifier "generally" must be emphasized before giving evidence of the validity of this premise. It is believed that a general negative attitude toward state and federal accident data currently exists, which is unjustified. (There is not a system in existence that cannot be criticized because it is not capable of doing something for which it was never designed. Current accident data systems are no exception. Rather than concentrating on the benefits that the data can produce, the focus is shifted to the vulnerable areas—things the systems cannot accomplish mainly because they were never designed for that purpose.)

Typical state accident records systems are now capable of problem identification, high accident location identification, and certain specific countermeasure evaluations. Increases in quality can bring about tremendous gains in these areas. However, as the data systems continue to be modified to produce greater data quality, these processing functions must not be put on hold pending better data that may never materialize.

Given this qualifier, the evidence supporting this section's premise is overwhelming. Is there a researcher who has not accessed the accident records data base to answer the simplest of questions and found that the needed variable was totally missing, not captured, not entered, or so misunderstood by the data collectors that its values were unusable? The unreasonable number of missing values and the gravitation to certain codes (e.g., inattention) are thorns in the side of those who are trying to turn data into information. Inconsistencies in the data abound in those variables, which are redundant in their potential information content.

These statements are based on first-hand experience with at least 10 different state records systems. Because there were no exceptions, these conclusions are not believed to be hasty generalizations. So, although there are some good data elements and a sufficient level of quality to generate some valid information for decision making, it will be a long time before the information benefits that optimal-quality state-level data are capable of rendering can be achieved.

If the consequences of inadequate data quality were limited to these problems, researchers could plot gleefully ahead toward greater quality. However, the problem tends to feed on the good data in an accelerating vicious cycle. Inadequate data quality in some data elements has led to limited application of all data. Legitimate applications are challenged due to known deficiencies. Most critical is the demoralizing effect on those charged with the use of the data. Why expend the effort to produce useful information from data known to be of inadequate quality? This question has too often led to deferring innovations until the data quality is improved, which is the key issue addressed here. Its counterproductive nature is discussed after the conclusions of the premises stated so far are presented.

CONCLUSIONS ABOUT PREMISES

Current Methods of Attaining Quality Are Not Sufficient

The following two approaches are used to improve the general quality of data obtained from state accident records systems: (a) appeals to altruism and (b) mandates. Because the quality of the data is inadequate, neither of these devices has been successful in motivating sufficient action. The following discussion concentrates on the data collectors as one source of the problem. However, other links along the data-to-information chain can be equally responsible for poor data quality. For example, poor forms design, inadequate edits, and flaws in data entry can be just as devastating to data quality as a neglect on the part of the data collector, who in most cases is a police officer.

Appeals to altruism are insufficient because competing activities may seem even more altruistic. The police officer will sacrifice a concern for detail (e.g., who was wearing what type of restraint by seating position) to remove the injured safely from the scene or to prevent a secondary accident by restoring traffic flow. In fact, the entire area of data collection is generally inconsistent with the paramilitary perspective that has been developed through years of training and experience. The police officer is trained to act, to take control of the situation, and to provide remedial and preventive effects by this action. Generally, police organizations are viewed by the public as being the custodians of expertise on traffic safety by virtue of their being on the scene. The idea that this experience can be coded into a computer and transferred to other organizations for information generation and dispensation represents a real loss of power in this regard.

Paramilitary organizations usually respond quite well toward mandates. To a large extent, the adequate data quality available in certain data elements is attributable to this factor. However, as the quality requirements increase, the officers begin to question why they need to provide this much detail and whether it is really necessary. Ineffectual data elements of unnecessary complexity can then become extremely counterproductive. Mandates are of little value here because they are impossible to enforce. In fact, they may bring a backlash of resistance if pushed too far.

Although police organizations have been used as the primary example, they are not intended to be a scapegoat. Data quality suffers in other parts of the system as well, and for the same reason, that is, current motivational methods are not working.

A New Approach to Quality Improvement Is Required

The only reason this conclusion might not follow from the previous one is the supposition that more of the same (i.e., more appeals to altruism or more mandates) might improve data quality. That these techniques have not produced the desired level of quality over the past 20 years is adequate evidence that a new approach is required. Instead of arguing this case further, an alternative called distributed responsibility is suggested. In this approach the data collectors (as well as the other weak links along the chain) assume respon-

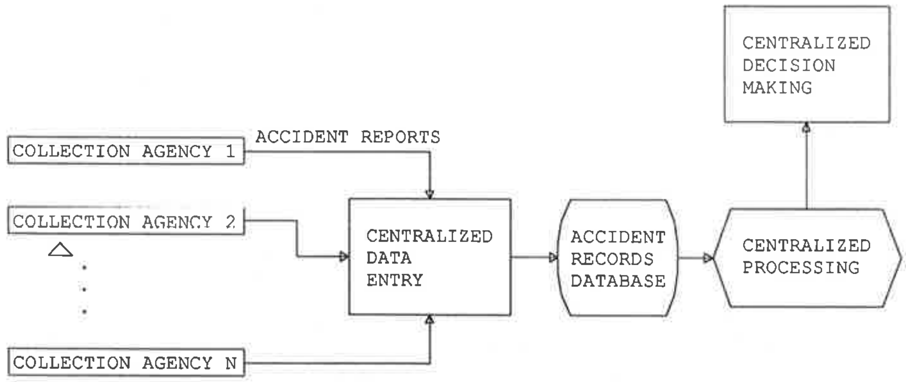


FIGURE 1 Traditional organization for centralized responsibility.

sibility for greater data quality. They do so because they find that it is in their personal or their organization's functional self-interest. Thus, the question of why so much detail is required is answered not by some mandating administrator but by the system itself.

A distributed responsibility system is data collector centered in that compellingly useful information is fed back to the organization that generated it. Although the examples presented focus on the law enforcement community, the applications are widespread. Because accident and citation information emanates from law enforcement, that community would be the target of these distributed responsibility benefits. However, in this era of integration, the same principles apply to medical data from hospitals, emergency medical data from emergency medical technicians, and even roadway characteristics information from engineers. The loop must be closed so that those responsible for data quality either get the full benefits of their labors or suffer the consequences of their deficiencies.

Figure 1 shows the traditional centralized responsibility model. In this paradigm each of the data collection agencies (e.g., state and local police) submit the data to a centralized data entry organization, after which the data go into the data base and are processed. The first attempt at distributed benefits is shown in Figure 2, where static hard-copy reports are sent back to the data collection agencies. (These agencies are collapsed into a single box to simplify the diagram.) Clearly, an alternative term for distributed responsibility is distributed benefits, because it is the benefits that will motivate the entire organization to assume proper responsibility. It is essential that the data collectors see the output and recognize its value. Traditional outputs—standardized hard-copy reports—have failed to provide this motivation because they have failed to obtain the participation of the data collection organization in their generation and use.

These two aspects of the process are inseparable: generation and use. Standardized hard-copy reports seldom answer the detailed questions that arise dynamically in the field. Rather, they tend to represent the only output of the process to the data collectors. Thus, if the specific required information is not in these static reports, the conclusion is that the entire process (data collection through output generation) is worthless, which further undermines the data quality for currently useful applications.

Distributed benefits must be technology based. These capabilities were not possible in the era of hard-copy outputs. Now, however, with the ubiquitous nature of personal computers, there is no reason that software to generate information useful to the local organizations cannot be given to them (costs are discussed in a subsequent section). This capability, coupled with the advent of new and advanced input technologies, can lead to an order of magnitude increase in

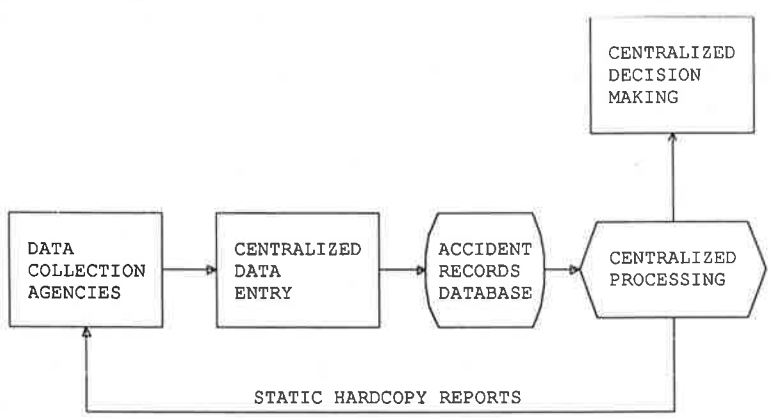


FIGURE 2 Initial attempts to provide distributed benefits.

the quality of data available throughout all organizations in the traffic safety community.

EXAMPLE IMPLEMENTATIONS OF DISTRIBUTED BENEFIT SYSTEMS

A number of distributed benefit systems exist, none of which has reached its full potential of user orientation and application. The purpose of this section is not to promote any given system or approach; rather, it is to promote the concept of distributed benefit systems in general. All of the systems mentioned, and the many others believed to be currently under development, should strive to provide local decision makers with information, not data. This information needs to be in a form that can be understood and used directly by the decision maker, recognizing that statisticians and computer technicians might not be available to local organizations.

A list of distributed systems is given in the Appendix. The list is not purported to be complete, and other systems will continue to be announced. However, these systems do represent a fairly good cross section of the current state-of-the-art. Table 1 presents a comparison of the various aspects of these systems as of March 1990, when several of them were presented in a NHTSA-sponsored conference (*1*). (Because these systems are in continual development, the representations are subject to change. The contacts given in the Appendix can be consulted for current information.)

Table 1 indicates that the current systems cover a wide range of data sources and entry responsibility. Most of the local systems require their own data entry. The demand for such systems indicates that a large number of them are already developed at local levels but have not been publicized. These systems will provide benefits by increasing data quality, especially for those data elements maintained at the local level. They will also promote awareness of the value and use of data that are sent to the state from local areas. A further benefit is the availability of the most current data to the local officials.

Figure 3 shows the locally autonomous systems. There are two independent data entry functions, which may be inefficient. Data entered at the local level are usually only a small subset of the total record. However, because the state and local data entry are redundant, there may be a problem of consistency between the two. There is also a question of whether the local concern for the data sent to the state is increased or decreased, because they may reason that their data needs are now satisfied. These problems are not insurmountable; the local data entry systems are just a stepping stone to that point in time when all data will be entered locally and uploaded to the state, thus eliminating redundancy and greatly increasing the quality of all data.

As indicated in the second entry of Table 1, a few states are providing a downloading service to furnish information capabilities to the local organizations. This model, as shown in Figure 4, is fully capable of returning a complete data access capability to the locality without redundant data entry. The model has the advantage of efficiency and consistency, but it is resisted by some jurisdictions because it suffers in data timeliness. However, because it gets a turnkey system into the hands of local officials in the quickest possible time with the largest possible data base, the timeliness shortcoming might be overlooked in the near term. Once the local data entry capability is established, the responsibility can be transferred locally (again, preferably in the direction of ultimate automatic uploading of locally entered records).

The third level of sophistication is with those states that already enter data locally and upload to the state level, as

TABLE 1 CLASSIFICATION OF DISTRIBUTED SYSTEMS

Classification	Systems
Autonomous: no interaction with state data base	MTRS
	SCARS
	KARS
	TARS
State-Dependent: Data downloaded from state	CARE
	TARS
Local-Entry–Dependent: Data uploaded to state	LANSER
	TRASER
Totally Local–Dependent	None currently known

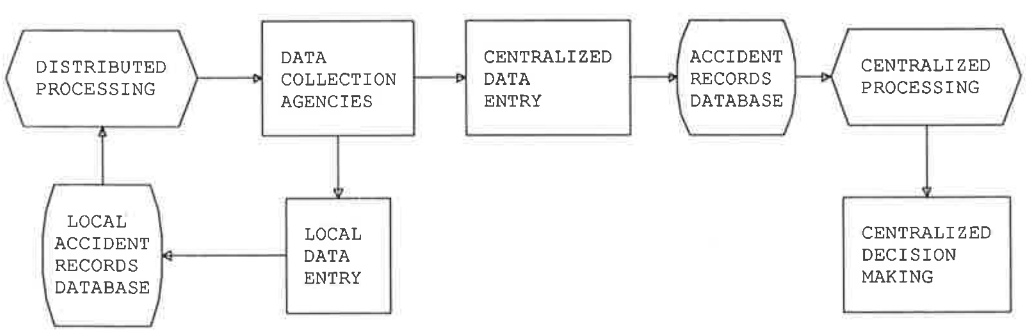


FIGURE 3 Locally developed and implemented microsystems.

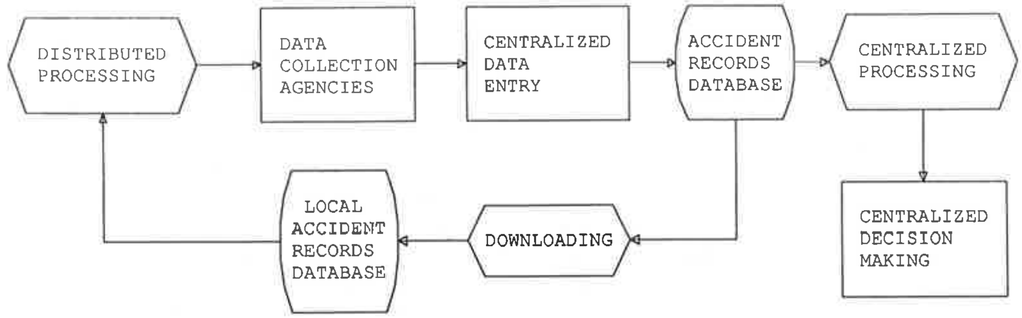


FIGURE 4 Distributed responsibility/benefits model.

shown in Figure 5. Here, networking capabilities are being exploited on a pilot basis, and the concept is proving to be feasible. Although no examples are given of a totally locally dependent system, a representative at the San Diego meeting (*1*) indicated that this was the case in Ohio. In this extreme, the local organizations own the data and develop their own processing capabilities. This method seemed satisfactory, but there clearly, is an optimal point of distribution of responsibility after which inconsistency between local jurisdictions can become counterproductive. All of these new distributed systems are still attempting to find that optimal point.

CURRENT DISTRIBUTED SYSTEM CAPABILITIES

The basic objective of the distributed microsystems is essentially the same: to provide local access to information for local decision makers who have little, if any, formal computer or statistical training. A further objective is to minimize the implementation expense. Therefore, these systems have been designed to operate on hardware that is largely available (mainly IBM PC/AT compatible). They all have the capability to create subsets of the accident records so that such types as alcohol, bicycle, and pedestrian can be examined. This capability ranges among the various systems in both sophistication and user friendliness.

Current network technology makes it feasible to enter all data from the local level. The most basic requirements for electronic communication between a remote site and the central data base include a modem and a telephone connection. The modem is a simple device that translates digital signals from a computer to and from analog signals carried by a telephone cable. With the addition of a dedicated phone line and modem, two remote computers can communicate easily.

Even when all police officers have personal lap-top computers in their squad cars, the process is not complete until the central data base is updated. Initially, this task will be accomplished by storing the data gathered by the officer onto a diskette, placing the diskette in the local office computer, dialing the central computer, and uploading the information. Information can be disseminated to local officials in the reverse fashion. Data collected and processed in the central location can be sent electronically over the telephone system to the remote sites.

A complete networking of all computers in a state's traffic control offices is an achievable goal. A centralized data base will then be available to all sites, and data-base updates will be made with no replication of data entry and no conflicting data. For a region's traffic data, a viable distributed system includes a local area network (LAN) with a central data base acting as a server for the remote clients seeking information. Data base updates will be available immediately to all servers on the network, and all servers will be able to modify the data base directly with their local input. Figure 6 shows such a distributed system. Although this model is currently not being implemented, the technology is currently available, and it is expected to be the direction that distributed state systems will take over the next decade.

The basic output of current systems is the frequency distribution for a given variable or set of variables. Some have standardized output report generators, which limit the users to predefined outputs. Although these outputs have little more than hardcopy flexibility, that they can be obtained for virtually any subset is a definite advantage over the model given in Figure 2. However, most have the capability, at least for

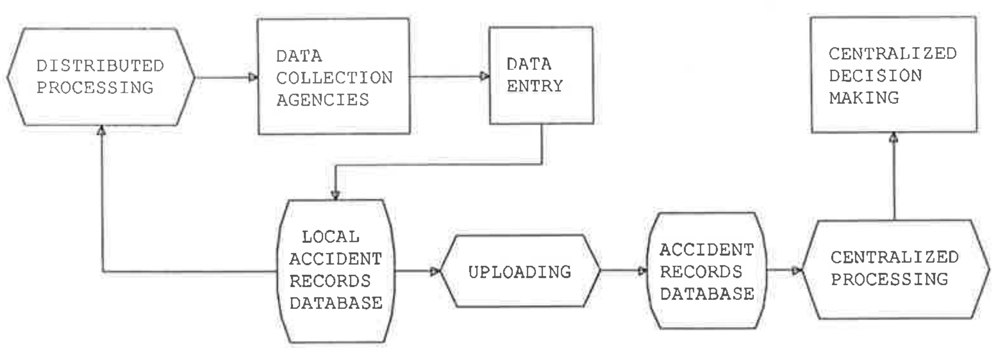


FIGURE 5 Distributed system under current technology.

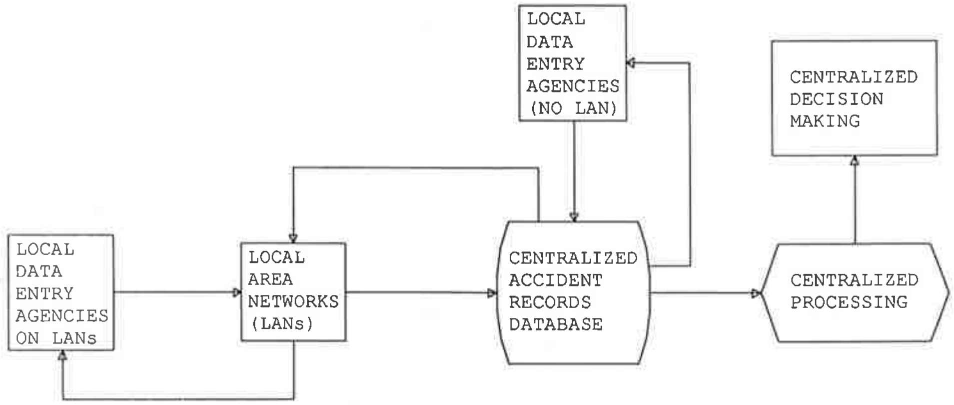


FIGURE 6 Networked distributed system.

the more advanced user, to bypass the standard reports and generate virtually any summary of information from the data base. Some are based on dBASE or can generate ASCII files so that their capabilities are extensible. The generation of any cross tabulations and the automatic identification of overrepresentations are two capabilities that are of major value.

High-accident locations can be generated easily on most of these systems using either user-specified criteria or the specification of an exact location. Additional information on a location can also be obtained by frequencies, cross tabulations, or, in some cases, standardized reports. For those wishing to retrieve the hard copy of the accident report, the accident numbers are readily available.

The use of graphical outputs is just beginning to find its way into local systems. Bar charts from frequency distributions and color within cross tabulations enable the users to visualize overrepresentations in a much more understandable way than viewing a list of numbers. Some of the systems can automatically generate collision diagrams, which seems to be feasible and should have widespread application within the next year or so.

As mentioned, the cost of a distributed system is minimal when the local departments are already equipped with computers, even if these are only IBM PC–compatible microcomputers that are being used primarily for word processing. Although some small police departments might not have this technology yet, it is expected that they will soon. With such equipment, the only hardware requirement is a centralized PC file server to store and protect the centralized data base, as well as modems at each remote PC. The server must contain a large (minimum 1.2-gigabyte) hard disk, and it must be accompanied by a backup system (possibly a tape drive costing approximately $6,500). The modems are required for data transmission over telephone lines, and the typical 9,600 bits/sec modem would cost about $400. If several PCs are employed in a larger department, it would be preferable to have them connected by a LAN. This would require the purchase of Ethernet cards for each PC (approximately $150 each) and server software, which can range in price up to $6,000.

Table 2 presents an approximate price list based on the establishment of a LAN consisting of five PCs at one locality. The costs in this table assume that the department effectively starts from scratch. Most local departments would not have this system dedicated to their accident records. They might be using it for departmental administration, citation processing, word processing, and a variety of other applications. Thus, only a small fraction of this cost would need to be justified by the accident records system. In many cases, this equipment is already installed for these purposes, and only a small marginal cost is required to install the accident-processing components. Such is also the case with a small department that already has a single PC with time and space

TABLE 2 POTENTIAL SYSTEM COSTS ($)

ITEM	UNIT PRICE	QUANTITY	TOTAL
386 PC & network card	2500	5	12,500
386 server	6500	1	6,500
Laser printer	3500	1	3,500
LAN software pkg.	6000	1	6,000
Modem	400	5	2,000
Coax cable (1500 ft)	500	1	500
Misc. Network Hardware	500	1	500
Labor			7,500
Training			3,500
TOTAL			42,500

availability and therefore only require a modem and a communications card.

CONCLUSIONS

The major conclusion obtained from considering these premises and their supporting facts is that innovations in technology applied to the entire data-collection–to–information process will create data quality, provided the human element is not neglected. The greatest technology advances will not render the data more usable if the data collectors and other support personnel are not properly motivated. The success of distributed responsibility depends on distributed benefits. It is essential that the data collectors and their respective local organizations be targeted for software designed to be used as an integral part of their jobs, which will give them a vested interest in seeing that the data are accurate and complete. In the meantime, prototype data should be used to continue innovation. Data inaccuracy should never be used as an excuse to delay innovation, for it is only by having the capability to generate benefits that the value of quality data can be appreciated.

ACKNOWLEDGMENTS

The research and findings presented in this paper were partly sponsored by the Alabama Highway Department, the Alabama ADECA Traffic Safety Section, and NHTSA.

APPENDIX
Contact List of Micro-Based Accident Systems

City Accidents Rapid Evaluation (CARE)
David B. Brown
Department of Computer Science and Engineering
Auburn University
107 Dunstan Hall
Auburn, Ala. 36849-5347
(205) 844-6314

Kansas Accident Records System (KARS)
University of Kansas Transportation Center
PC-Transmission
2011 Learned Hall
Lawrence, Kans. 66045

Local Area Network Safety Evaluation and Reporting System (LANSER) and Traffic Services Microcomputer System (TRASER)
Marlin Crouse and Barbara Delucia
Texas Transportation Institute
Texas A&M University System
College Station, Tex. 77843-3135

Small Computer Accident Records System (SCARS)
Ken Courage
McTrans Center
University of Florida
512 Weil Hall
Gainesville, Fla. 32611

Traffic Accident Reporting System (TARS)
Steve Lau
Lau Engineering, Inc.
17220 Newhope Street, Suite 204
Fountain Valley, Calif. 92708-8771
(714) 546-2046

Virginia Micro Traffic Records System (MTRS)
Robert Breitenbach and Jill F. Davis
Transportation Safety Training Center
Department of Justice and Risk Administration
816 West Franklin Street, Box 2017
Richmond, Va. 23284-2017
(804) 367-6235

REFERENCE

1. *Proc., Conference on the Collection and Analysis of State Highway Safety Data*, San Diego, Calif., NHTSA, U.S. Department of Transportation, June 1990.

The opinions expressed in this paper are those of the authors and do not necessarily reflect positions held by the sponsors.

Publication of this paper sponsored by Committee on Traffic Records and Accident Analysis.

Intelligent Vehicle-Highway System Safety: Approaches for Driver Warning and Copilot Devices

ANTHONY HITCHCOCK

Estimates of accident savings ascribed to the intelligent vehicle-highway system (IVHS) and IVHS devices have been published by Mobility 2000. These estimates were obtained through the use of expert judgment applied to various configurations of accidents, for which data are available in standard sources. Similar work has been done in Europe. The availability of in-depth data sets and other more detailed sources enabled estimates to be made on a quantitative basis. Methods used in European and U.S. work are discussed, and results are compared. An account is given of a possible development of the evaluation technique that can provide greater precision and might facilitate the choice between different realizations of the same requirement. Finally, the possibility of interactions among different devices on the same or different vehicles that might hinder safety is identified. The need for systematic configuration management of IVHS is emphasized.

Techniques developed in Europe for the examination of intelligent vehicle-highway system (IVHS) safety are described. The use and extension of these systems in U.S. conditions are discussed. The lack of certain data is identified as a barrier. Specific topics covered include the following:

- Reevaluation of the conclusions of the Mobility 2000 evaluation of IVHS (*1*) in light of some European results and discussion of the availability and collection of suitable data in the United States;
- Proposals to extend the methodology to enable quantitative evaluation of specific designs; and
- A discussion of the interaction between faults and of the need for an appropriate configuration management technique.

COST-BENEFIT ANALYSIS OF IVHS SAFETY

Accidents Avoided by New Technology—U.S. and European Approaches

The Mobility 2000 group desires to bring together practitioners involved in the development and exploitation of the IVHS. The group recently issued a series of reviews and an initial cost-benefit analysis of IVHS (*1*), which addressed the impact of the introduction of the IVHS on highway deaths and injuries. Several possible IVHS devices were hypothesized, and five classes of accident were identified as susceptible to change. Estimates of the effect on each were made on the basis of expert opinion. The five classes were related to relative motion at the time of collision: off-road, head-on, rear-end, angle, and sideswipe. In forming their judgments, the experts clearly had to picture the mental processes by which drivers make the errors that lead to such accidents. They then had to judge to what extent a warning (or preemptive control action) would be beneficial.

This approach raises a number of questions:

- Have all the routes by which accidents of each type can occur been considered or identified?
- Is the approach not highly subjective? Is it reproducible?
- Is the approach not open to accusations of bias?
- Might the approach not be biased by the choice of accident classification (which is constrained by the form of the existing statistics, chosen for other purposes)?

These questions can be robustly answered by the observations that work must begin somewhere, that it is more important to obtain answers correct to order of magnitude now than pedantic precision next year, and so on. Indeed, the questions are necessary and timely; because expert opinion will be the immediate source of decision making, it is best that it be expressed.

However, the Mobility 2000 authors were apparently unaware that a quantitative method for tackling this problem had been employed by Hitchcock (*2*) to provide a similar preliminary estimate of the potential impact of the European IVHS project, PROMETHEUS. An account of this project has been published (*3*). The technique requires fewer heroic subjective assumptions, although it is not free from subjective elements. Further papers on this technique have been published by Broughton (*4*) and by Fontaine et al. (*5,6*). The latter developed the technique usefully.

These sources refer to the types of device being developed in PROMETHEUS, which the Mobility 2000 reports call the advanced vehicle control system (AVCS). AVCS-1 devices fall short of full automation and are therefore applicable on all roads, both in urban and in rural areas. [AVCS-2 involves full automation of the driving task on certain freeway links, whereas AVCS-3 adds strategic routing and scheduling functions to an automated freeway network (*1*).] Discussion here is largely restricted to AVCS-1 devices.

Hitchcock (*2*) used the U.K. at-the-scene data base, which consists of observations made by professional observers at the

Program on Advanced Technology for the Highway, Institute of Transportation Studies, University of California (Berkeley), 1301 South 46th Street, Building 452, Richmond, Calif. 94804.

scene of accidents within half an hour of their occurrence, supplemented by interviews with those participants in the accidents who consented to them. A tolerable response rate was achieved, with full sets of interviews being conducted in about 60 percent of all cases. Some 1,300 full records were used. Hitchcock also gave a rather general description of the way in which PROMETHEUS would achieve unprecedented safety, placing insufficient emphasis on advice to drivers as opposed to direct control action.

Each relevant accident record was read, and a judgment was formed about whether, in the circumstances described, the IVHS devices specified could have affected the course of the accident. By totaling the number of successes for each device, an upper bound to the efficiency of the devices can be obtained, after making some corrections arising from the nature of the sample.

Broughton (*4*) used the same data, but provided much more precise descriptions of the functions offered by possible PROMETHEUS devices. He also employed less restrictive assumptions than did Hitchcock, who had been concerned that IVHS devices should not perform in ways that led them to be mocked, such as stopping at green lights to avoid collision with traffic approaching a red or declining to leave the road to enter an owner's driveway.

Fontaine et al. (*5,6*) also gave more precise descriptions of 14 possible PROMETHEUS functions, but, lacking an in-depth data base, used the precise police reports of accidents known in France as *procès-verbaux*. These authors appreciated better than their predecessors the importance of precisely stating the assumptions made about the scope of action open to the device, thus avoiding unexplained differences such as those between Hitchcock and Broughton. Fontaine et al. also describe an alternative method for typifying accidents, but it does not add to the quantitative results.

Conditions are different between Britain and France, but they are probably more similar than either is to the United States. At full penetration of the devices, Hitchcock found an upper bound to the reduction in accidents to be approximately 25 percent, Broughton found 60 percent, and Fontaine et al. found 45 percent. Each recognized that the figure was an upper bound on the basis of the following assumptions (*4*):

- All devices work as intended;
- All drivers react as intended; and
- There are no unintended side effects, particularly no behavioral changes.

The devices assumed by these authors are not the same, however. In particular, Broughton assumed a device that could advise the driver of the speed "most appropriate to the conditions," varying with congestion, traffic pattern, weather, light, and so on. It is not clear that anyone could arrive at a consensus, before the accident, on this speed. It is even less clear that the speed could be specified mechanistically. Hence, drivers might be more reluctant to accept its guidance than that of other copilot devices. Without this assumption, Broughton's figure would be approximately 20 percent less—more in line with that of Fontaine et al. whose less optimistic speed control has an effect in only 3 percent of cases.

Fontaine et al. and Broughton find that nearly half of the remaining effect (i.e., 20 percent of all cases) can be ascribed to devices in which active elements in two vehicles determine the likelihood of collision by communicating their positions and intended trajectories. Such devices are clearly not effective until penetration of the market is nearly complete. Hitchcock did consider such devices, but, under his more restrictive assumptions, its effect was only 5 percent.

These large corrections bring the three results into line to an unjustified extent. There are other smaller differences that, if corrected, would take them further apart. Nevertheless, the three European studies are much less at variance than appears at first sight. It seems right, at full penetration, to reject Hitchcock's restrictive approach to intervehicle communication. It seems best, too, to reject Broughton's so-called "wise" speed indicator. The conclusion, then, is that the effect under European conditions of PROMETHEUS is a reduction of 45 percent or so in injury accidents.

There is some agreement that a cooperative trajectory-indicating anticollision device—perhaps as advanced an application as considered—would reduce accidents by only 20 to 25 percent. The U.S. workers also considered an advanced robust collision prevention system, which reduced all categories of accident, at full penetration, by 70 percent. This variation may reflect optimism in the United States or actual differences between U.S. and European conditions.

Although all authors, European and American, recognize that it is possible to mount devices in a vehicle that can detect the presence of alcohol, none listed such a device. Those who do include an impairment detector based on erratic behavior consider that it will not detect alcohol use.

Application to U.S. Conditions

Although it is relevant to draw attention to large discrepancies (as with the robust collision prevention system mentioned previously), little consideration is needed to demonstrate that European numbers are not transferable to U.S. conditions. In the United Kingdom, for example, roughly 15 percent of vehicle-miles traveled (VMT) is on freeways. In California, by contrast, some 40 percent of VMT is on freeways. Equally, in both Britain and France about 40 percent of those killed (to compare numbers injured is unwise because of differences in definitions) are either pedestrians or pedal cyclists. The corresponding figure for California is 20 percent.

Hitchcock (*2*) indicates that the IVHS devices he considers will avoid a larger fraction of freeway accidents than of all accidents. No doubt, other, subtler factors related to differences in terrain, law, and social custom are involved, but it would not be useful to pursue them here.

It is not immediately possible to apply the techniques directly to U.S. conditions because of a lack of necessary data. In Britain an in-depth data base was used. A publication by the Organization for Economic Cooperation and Development (OECD) (*7*), reviewing data bases of this type, refers to only two such data bases in North America, although its time scales preclude reference to the pioneering Indiana trilevel data (*8*), which are of the required type. The Indiana data base dates from the early 1970s and therefore may lack the power to enforce conviction.

One data base referred to by OECD (*7*) is the Canadian work done along with the trilevel work, which was concluded in 1979 and therefore may suffer the same lack of power to convince. The other is the National Accident Sampling System (NASS) (*9*).

NASS data are detailed but emphasize the crash and post-crash phases of an accident, as expected, given the primary responsibility of its sponsor for vehicle safety standards. In the form in which NASS data are available, their value in this context may be limited. Work is in progress to determine to what extent they can be used.

Equally, there is no U.S. equivalent of the *procès-verbal.* Although U.S. police records of accident investigations were not examined, they are said to be extremely variable. Indeed, that the levels of emphasis of good reporting vary from police force to police force and, no doubt, from officer to officer is a matter of record.

However, there are also records that are complete. There is no a priori reason to suppose that the quality of the police record (regarded as a source for research of this kind but not written for that purpose) is closely correlated with any other relevant characteristic of the accidents. Tolerably quantitative results could possibly be obtained by using a selection of existing police records.

There are obstacles to this approach. Records that identify individuals, their statements to police, and police opinion for situations in which prosecution is possible are rightly regarded as confidential. The police would carefully examine the protocols proposed by any researcher before allowing access to such data.

There may, however, be ways of avoiding this difficulty, at costs in time and money. Perhaps someone who was permitted to view such data could sanitize it by removing names, addresses, and the like, passing only extracts to the researcher, or could arrange for letters to be sent to those involved seeking permission for data to be passed after being made anonymous. If the number of records to be examined were large, such expedients would be costly. However, if the researcher were content to examine 1,000 or so accidents, which would certainly suffice for the first stage, the cost would not be prohibitive.

Another possibility would be to work in an area where the police could be persuaded to make their records more complete for a limited period. Unfortunately, it would not take long to assemble records of 1,000 accidents in a small area.

It is unlikely that a fully representative sample of accidents could be assembled by such methods, but the experiment of assembling data could be useful. (The representativeness of the European data can also be questioned.) It would perhaps be appropriate, in view of the European findings, to take separate samples of freeway and nonfreeway accidents.

EFFECTIVENESS OF ALTERNATIVE DESIGNS

The evaluation technique can also be developed for use at later stages of design. It can probably be modified to select the more efficient alternative realization of a specification and to examine the significance of engineering tolerances in the specification. The basic tool required is a computer simulation shell that can represent the movements of several vehicles on a road or on roads of variable geometry. It must be possible to insert varying control algorithms. In such a simulation the human driver must be represented as a controller whose objective function is a particular space-time trajectory (initially the one actually followed). The driver is represented as being able to change the objective function on receipt of an external stimulus (personal observation or IVHS-generated warning) with a variable time lag.

The computer shell is then used to simulate a series of actual accidents, the course of which is known either because the accidents are recorded in an in-depth data base or in some other way, as discussed previously. As in the earlier case, it is assumed that one or more IVHS devices, including the one being tested, are installed in one or more of the vehicles in the simulation. One or more alternative assumptions are entered about how the driver's control-objective trajectory will change if a warning is received. These assumptions must be consistent with what is known about the way in which the driver actually behaved. Alternatively, if the IVHS device is one that overrides the driver, the user must feed in the device's behavior.

In practice, a series of variations of a single accident case would be considered, varying driver reaction times, precise initial positions, and speeds of vehicles. It will always be necessary to check that, without warning, the accident occurs as it did. It can then be determined whether the effect of the device under evaluation is always the same. If desired, the behavior of the IVHS device can be varied within its engineering tolerances. This process must of course be done for a series of accidents taken from the data bank in which there is reason to suppose that there will be an effect.

The development of a computer simulation of this kind will be lengthy and complex, and its use in the way described presupposes not only the existence of a significant number of in-depth accident reports but also significant expenditure in setting up each case. Fairly good statistics are required, too—at least 20 and perhaps as many as 100 relevant cases should be run.

However, the proposed method does ensure that design decisions are tested against real data. Experience in the direct use of in-depth data bases reveals how much more complicated most accidents are than appears in a simple statistical record. Vehicle movements and drivers' intentions are affected by the presence of road users and vehicles who are not involved in a crash, have committed no irregular action, and may not have been aware that anything was amiss.

The alternative to this kind of analysis is to design the device to meet simple, common-sense criteria. These criteria are likely to be based on a simple mental model that classifies accidents into a small number of categories and assumes uniformity within each category. However, if the designer believes that the requirement specification is self-justifying, the error will not be detected until application. The remedy for oversimplification of a problem is exposure to fact. No other way of achieving this than the expensive suggestion presented earlier has yet been proposed. People are complicated—there is more than one way to make a mistake.

INTERACTIONS BETWEEN FAILURES

Most accidents occur after several distinct driver errors or other system failures. However, if all bad driving led to an accident, the roads would be much more dangerous than they are. In the future, it must be expected that most faults in IVHS devices will be detected and corrected before an accident occurs, partly because of the safety-oriented reaction of other drivers and IVHS devices to hazardous behavior caused by a faulty component. A potentially hazardous fault in a component may be cushioned by corrective behavior in the surrounding traffic.

Alternatively, corrective behavior may not be necessary because the IVHS design is such that, although the affected vehicle or road device has lost a capability, no immediate hazard arises. However, when two or more such cushioning groups meet or a diminished-capability vehicle comes across another failure, the level of hazard may increase. If there are active automatic controls (as opposed to warnings) or if false information is being propagated and relied on, the hazard level may exceed the level that would exist in the absence of IVHS devices.

Freak Waves

Freak waves are a known phenomenon of the oceans, occurring well away from coasts. The irregular movement of the surface of the ocean is a superposition of many separate wave trains of varying amplitude, direction, and phase. Most of the time, interference between waves keeps the mean displacement relatively small, but constructive interference occasionally occurs. Then, a freak wave, 10 or 20 times the mean amplitude, appears suddenly out of a calm sea. Ships have been wrecked by such waves.

The phenomenon can arise because ocean wave trains are dissipated very slowly; hence, a train can readily travel thousands of miles, in deep water, from the storm that generated it.

Could a similar phenomenon occur in traffic under an AVCS? Waves can certainly travel through traffic streams. The shockwave phenomenon, by which a disturbance in a traffic stream is propagated backward down the road, is well documented both in car-following theory and by observation. A speed disturbance grows as it propagates from front to rear of a close-spaced platoon, unless special care is taken to include stabilizing terms in the control function to damp the waves (*10*). Even without a positive attempt to maintain a fixed distance from the preceding vehicle, some undamped waves could be propagated through a traffic stream if there are devices present that attempt to avoid a close approach to vehicles ahead and to either side.

Advanced AVCS-1 systems may include devices that communicate from vehicle to vehicle, potentially over some distance. These systems might also be designed to cause an automated avoidance response. Again, the opportunity for creation of an undamped wave could arise, although it is not possible to be precise until the intended function of the system is evident.

The freak wave analogy can perhaps be carried too far. It is mentioned merely to suggest one mechanism by which a multivehicle interaction can produce a hazardous situation, so that the investigation of interactions does not seem pointless. In the search for deleterious interactions, however, the researcher should consider if intervehicle communication and subsequent actions can produce waves that would be propagated along or across a highway.

Hazardous interactions might also occur among different devices on the same vehicle, although in this case the freak wave analogy would not apply.

It is relatively easy to postulate credible pairs of IVHS devices that could interact in a hazardous way, even without faults. One vehicle might have a lane-keeping device that, when there is another vehicle alongside, keeps it a distance x from it. Another could have an anticollision device that causes it to veer away if another vehicle comes within a distance y of its side. If $x < y$, there will be a hazardous interaction.

Another possibility is that two such lane-keeping devices, with compatible steady-state objectives, could interact to produce divergent oscillations because their control functions had been designed independently (which could happen even with two different devices from the same manufacturer). More complex, more likely, but less readily imaginable combinations no doubt also exist.

Configuration Management

Techniques for looking for interactions among AVCS devices are not discussed. Part of the data for such an analysis are clearly specifications of the devices concerned.

Some devices currently being produced will still be in service 25 years hence. If a new device is designed at that time and the designer wants to search for interactions, he will have to research the devices placed on cars in 1991. It does not seem likely that these data are being retained. Indeed, if an engineer today wished to carry out an analysis of the safety of existing IVHS devices, the engineer may find that some of the necessary specifications have already been lost as staff changes and record holdings are reduced.

In other fields involving complex systems, such as spacecraft or aircraft, a technique known as configuration management is used to maintain comprehensive records of interacting components. The system monitors whether individuals concerned with one component have been informed about changes to the other component; modifications, updates, and improvements to hardware and software; repeat analyses of the safety-critical part of the system following such changes; and so on. In these cases, of course, the whole system is the responsibility of one body, which is in a position to exercise such control and has an interest in doing so. Otherwise, if the system doesn't work, the reason cannot be determined. In IVHS the legal situation is different. The components have various owners, and not all of them have any interest or responsibility for system safety. But the need for the ability to determine the configuration is no less. The owners expect that someone has accepted responsibility for determining that the system as a whole is not excessively hazardous. They believe they have a right to be assured that any proposed addition to the system does not make it more hazardous. The courts are likely to identify those "individuals" after an accident, if no one does so earlier.

No attempt is made here to suggest where responsibility should lie or by what procedures it should be discharged. However, this task will not become easier as time passes.

Specifications

A first step in rectifying these deficiencies is to determine what the content of a formal specification of an IVHS device should include. What needs to be recorded so that analyses of safety-critical systems containing the device will be possible and valid? This task will require additional analyses, as well as a good deal of communication among the involved groups.

It is suggested that entries will be necessary under the following headings:

1. *Initiation.* What stimulus causes the device to function?
2. *Modes of Operation.* Under what conditions (e.g., vehicle in motion, crashing, traffic lights red, official emergency) does the subsystem function and under what is it inhibited?
3. *Effect.* What does the device do?
4. *Changes.* What can the device modify externally (e.g., heading of vehicle or aspect of traffic light)?
5. *Outputs.* Does the device signal that it has completed its task? Does it give intermediate signals? Are there any fault signals?
6. *Faults.* How does the device behave in foreseen fault conditions?

CONCLUSIONS

A number of suggestions have been made. It is hoped that the data needs identified will be met, that it will be possible to demonstrate that quantitative approaches to the evaluation of IVHS safety are possible, and that the use of these systems will result in increased safety.

The following are suggested:

- To collect a small set of in-depth accident reports applicable to U.S. conditions, would be useful especially for freeway accidents, for use in these kinds of analyses.
- Advances in determination of the relative effectiveness, from a safety viewpoint, of different realizations of the same IVHS requirement specification are possible by constructing suitable shells for computer models.
- As IVHS devices multiply, the possibility of interactions among them will loom larger. Thus, full functional specifications for all IVHS devices present on the roads must be available to analysts in a standard form. A first step would be to initiate consultations about standardization, with the objective being to define effective configuration management protocols. The alternative approach—waiting for accidents to occur—may cost lives and is likely to be expensive for manufacturers and highway authorities.

ACKNOWLEDGMENTS

The research reported here is part of the Program on Advanced Technology for the Highway (PATH), prepared under the sponsorship of the State of California Business Transportation and Housing Agency, Department of Transportation, and the U.S. Department of Transportation, FHWA and NHTSA. The author would like to thank his colleagues at the University of California, Berkeley, particularly Susan Graham, Robert Parsons, Steven Shladover, and Wei-Bin Zhang, for helpful discussion. Appreciation is also extended to Anna Bozzini and Catherine Cortelyou for invaluable bibliographic and other assistance.

REFERENCES

1. *Final Report of the Working Group on Operational Benefits.* Mobility 2000, Dallas, Tex., 1990.
2. A. Hitchcock. *Potential Safety Implications of the PROMETHEUS Project.* Report WP/S&T/3. Transport and Road Research Laboratory, Crowthorne, England, 1987.
3. A. Hitchcock. Road User Safety: Possible European Research Cooperation. In *Roads and Traffic Safety on Two Continents, Gothenburg 1987.* VTI 328A. Linköping, Sweden, 1988, pp. 187–201.
4. J. Broughton. *The Possible Effects of Future Technological Developments on Road Accidents in Great Britain.* Report WP/RS/80. Transport and Road Research Laboratory, Crowthorne, England, 1988.
5. H. Fontaine, G. Malaterre, and P. van Elsande. *Evaluation de l'Efficacité des Aides à la Conduite.* Rapport INRETS 85. Paris, France, 1989.
6. H. Fontaine, G. Malaterre, and P. van Elsande. Evaluation of the Potential Efficiency of Driving Aids. In *Proc., Vehicle Navigation and Information Systems Conference*, Transport Canada, 1989, pp. 454–459.
7. *Road Accidents: On-Site Investigations.* Organization for Economic Cooperation and Development, Paris, France, 1988.
8. *Tri-Level Study of the Causes of Traffic Accidents.* Institute for Research on Public Safety, Indiana University, Bloomington, 1973.
9. J. O'Day, A. Wolfe, and R. Kaplan. *Design for NASS.* Highway Research Institute, University of Michigan, Ann Arbor, 1976.
10. S. Shladover. *Operation of Automated Guideway Transit Vehicles in Dynamically Reconfigured Platoons.* Report UMTA-MA-06-0085-79-1, 2, and 3. UMTA, U.S. Department of Transportation, 1979.

Publication of this paper sponsored by Committee on Vehicle User Characteristics.

Intelligent Vehicle-Highway System Safety: Problems of Requirement Specification and Hazard Analysis

ANTHONY HITCHCOCK

When accidents occur on a fully automated freeway, driver error will rarely be relevant. Design should not err in ensuring that equipment failure will not cause hazards. An attempt to demonstrate satisfactory methods of requirement specification and hazard analysis is described. The case treated is a single automated lane on a freeway that also has lanes for other vehicles and shares on- and offramps with them. When hazards are specified, it appears that the system configuration is determined. A number of concepts have emerged and seem likely to be basic to all hazard-avoiding designs. These concepts are identified and described. No method has been found for demonstrating that the set of hazards is complete. Peer criticism of those proposed is therefore earnestly sought because, as explained, these are the axioms on which the logical structure that demonstrates safety is to be based.

One form of the intelligent vehicle-highway system (IVHS) described by Mobility 2000 (*1*) is the advanced vehicle control system (AVCS). In its more advanced forms (AVCS-2 and AVCS-3), this term refers to systems in which, for a journey on a freeway, the driver of a vehicle surrenders control to a fully automated system, possibly retaining control of the route or merely specifying a destination. AVCS-2 and AVCS-3 are not seen as feasible options for off-freeway traffic; it would be neither possible nor desirable to automate the movements of pedestrians or other local traffic, and it does not seem feasible to create a hazard-free system containing such elements.

The proposed method begins by considering factors that can go wrong with a system and evaluating the consequences, together with an estimate of their probability. It will be worthwhile to attempt to design out some of these possibilities, whereas others are not sufficiently likely. By means of this preliminary hazard analysis, hazards of different degrees of seriousness are defined. Provided nothing has been overlooked, a safe system will result if the design is such that the hazards do not arise even if the system fails. This method, known as hazard and operational analysis (HAZOP or HAZAN) (*2*), is similar to one widely used in the chemical, nuclear, petroleum, and aerospace industries.

For fully automated freeways, there is no current experience from which probabilities can be assessed. However, there is obviously no way of guaranteeing that a car will not fail in operation, particularly if the owner is responsible for maintenance. If a vehicle fails in such a way that it decelerates more rapidly than a following vehicle can by braking, there is a clear danger of collision. Shladover (*3*) has shown that, in such a case, a collision can be avoided only if vehicle spacings are much greater than normal. Shladover also states that, if the vehicles are very close together (say 1 m or less), the collision only involves a small velocity change (delta-V) and is likely to result in property damage only. From this assumption comes the concept of widely separated platoons, in which vehicles are separated either by a very small or a very large spacing.

It is postulated that catastrophic hazards are those system states that are precursors of high–delta-V collisions, whereas critical hazards are the precursors of low–delta-V hazards. The design should exclude the former, even if a vehicle or some other part of the system fails or is circumvented deliberately. The incidence of critical hazards should be reduced as far as reasonably possible.

A set of hazards has been formulated. There does not seem to be any systematic method of demonstrating that the list is complete, although it is reasonably simple to demonstrate that a proposed condition is or is not a hazard. The list is therefore presented for peer criticism.

It is now permissible to say that a system is safe if it can be proven to avoid the defined hazards (except perhaps in some known cases of low probability or an excessive number of coincidental simultaneous failures). If the system is fully defined, the question of whether it is safe can in principle be answered definitively as a result of a process of mathematical logic. If a system can be shown to be safe in so unambiguous a fashion, it is clearly desirable to do so. There is thus a need to verify the definitions of hazards, which have the same relation to this analysis as do Euclid's axioms to his theorems.

One possible system that will be subjected to such analyses is described. To avoid the hazards that have been defined, the basic configuration seems to be constrained to a handful of possibilities. Further, some concepts have emerged that seem likely to be common to all or many such systems. Although the hazard analysis has barely begun (and is not reported), it is worth recording these results.

In the system considered, there is one automated lane in each direction, and that lane operates at a capacity of at least twice that of a nonautomatic lane. (If the system is to be economic, its capacity must be greater than that of a nonautomatic one.) If there is a network of such lanes, changes between links are not made automatically. This assumption

Program on Advanced Technology for the Highway, Institute of Transportation Studies, University of California (Berkeley), 1301 South 46th Street, Building 452, Richmond, Calif. 94804.

seems reasonable for the initial stages of exploitation of the technology. In Mobility 2000 terminology (*1*), this system is an AVCS-2 system.

APPROACH TO HAZARD ANALYSIS

Modular Structure

An AVCS system consists of a large number of separate, distributed control systems that interact with each other over time. Some of these systems contain hardware and produce mechanical output—for example, the control system that controls the steering axle. Others, like the program segment in a roadside controller that admits a vehicle, are all software in that their output is information sent to other parts of the system. It seems useful in the current context, however, to treat this difference as of little importance and to regard all modules simply as modules, with precursors, effects, and stimuli to other modules, possibly at later times, without regard to the electronic, electric, or mechanical nature of the means by which the effects are achieved.

This modular property of the requirement specification is vital to freedom from hazard in practice. Different manufacturers will achieve the desired effects in different ways, and, as time passes, many improvements will be made to each module. Modular construction of the specification eliminates the need to reassess the safety of the whole system every time a change is proposed.

Thus, it is both necessary and possible to regard the system as an assembly of modules, each of which must be fully specified insofar as its precursors, effects, and consequents are concerned and which must in total add up to a complete specification. A complete specification means that, for any conceivable configuration, either the specification will indicate unambiguously what happens next or it will be possible to demonstrate by deductive logic that no possible combination of circumstances could have preceded that configuration. This requirement is formidable, and it is not discussed further here.

The modular approach described implies an analogy between a distributed system and a software suite, the modules being paralleled by software procedures. This approach, in turn, opens to application the body of techniques for safety-critical software described in the literature (*4*).

Safety-Critical Subsystem

Some parts of an IVHS system are essentially concerned with actions that are critical to safety; others are not. A design principle recommended in relevant standards (*5,6*) and adopted here is that the safety-critical subsystem be kept as small as possible and rigorously separated from the rest, so that interaction between the two occurs only at clearly defined modules.

The focus here is not on the non-safety-critical part of the system, which does not mean that it is peripheral to the system as a whole. The entire subsystem concerned with controlling the movement of vehicles so that capacity is maximal (i.e., link control) can, and should, be kept separate from the safety-critical elements. So should the process (i.e., strategic control) by which drivers indicate where they want to get off and are given information about any parts of the system that are not functioning properly. These processes are vital to the proper functioning of the system but must be made peripheral to the hazard analysis by appropriate specification of the requirement.

One component of the system that is not modular and whose status as part of the safety-critical subsystem can be argued is the relevant law. Law is certainly part of the system, for the mechanisms may not prevent what the law admits as permissible. Thus, the designer must specify legal and administrative requirements along with software procedures or bounding conditions for stability of a control system. If an action violates the law, it may be hoped that its frequency will be reduced; however, it would be foolish to assume that the action will not occur. Indeed, the system should be designed, so far as possible, to make it difficult (as well as detectable and provable beyond reasonable doubt) to indulge in dangerous joy riding or otherwise to use the system in a way that endangers others.

There are possibilities of more organized wrongdoing. Someone could think it fun to sabotage the system—for example, by causing all police cars to run round in ever-diminishing circles or to track the movements of an erring spouse. Others may see the traffic control system of the wealthiest cities in the world as a natural target for terror. The system should be designed to protect against these possibilities, too.

It is assumed that the law will require the following, but it is not assumed that all drivers will conform:

1. Every vehicle entering the automatic system must have a license certifying that control systems conforming to a legally defined specification were present and working on some date past and that they have not been subsequently invalidated for some wrongdoing. (The period of the license is not critical to this discussion.) The control systems must be present and functional within defined tolerances at the time of attempted entry. It will be an offense to carry equipment on the vehicle that can issue signals similar to those specified for the system. (There will need to be a body of law and procedure for approving new and improved control devices.)
2. It will be proper for the license to be invalidated mechanically, with the possibility of appeal to a human court (and, presumably, financial compensation if the automatic checking device made an error).
3. It will be an offense to fail to be ready to resume manual control on reaching the designated point of departure or the end of the system. There will be a storage place, known as the dormitory (explained later) at the extreme downstream end or at convenient intermediate points. (It might be urged that those who end up in the dormitory create a rebuttable presumption that they have more than 80 mg or 100 ml of ethanol in the bloodstream, but this element is not a safety-critical one.)
4. It will be an offense to enter the automated system while carrying an external load or with a trailer that does not have appropriate control and communication equipment.

5. Devices that will be carried by emergency vehicles to permit abnormal movements may not be used on other vehicles, nor may their signals be forged.

Considerable controversy about the evidential nature of the system records in the absence of human corroboration will likely take place.

Definitions

It is assumed that the following vehicle controls will maintain the platoon formation:

1. A lateral control system will keep the vehicle in its lane by means of a passive reference line in the road. The reference can be coded to give data that will reduce lateral jerk by warning of changes in curvature and also to indicate the positions of off-turns.
2. A longitudinal control system will use an active sensor to determine the distance to the vehicle ahead and its time derivatives. The range at which the sensor can detect a vehicle ahead may be limited, either because it will not be able to detect a vehicle around a horizontal or vertical curve or by its nature. The maximum distance at which the sensor can certainly detect an object ahead is defined as the sensor-range spacing. This distance may depend on the weather. The system supervisors may choose to select a shorter spacing.
3. A vehicle-to-roadside communication system will be included. It will be of short range so that the noise level at any point ascribed to the presence of other communications (e.g., from parallel lanes or other vehicles at some distance) is small.

The first two controls are chosen in the light of current developments. The idea that communication ranges should be short to reduce noise is the author's. If this feature turns out to be critical, it should be checked.

It is also useful to define the following quantities, all of which are weather-dependent and most of which depend on the speed of the vehicle concerned:

1. Full platoon braking is the maximum deceleration that a platoon can reasonably be expected to undergo without collisions or a break in formation. (If each vehicle in a platoon brakes at its maximum rate, the rates will differ, and there will be low–delta-V collisions.)
2. Platoon spacing is the distance that must separate two platoons if the leading one suffers a deceleration considered to be the worst likely in case of catastrophe, the other simultaneously undergoes full platoon braking, and there is to be no collision between platoons.
3. Sensor-range spacing is the maximum distance at which the sensor can detect an object ahead (as defined previously).
4. Sensor-range speed is the greatest speed from which a vehicle can come to rest within sensor-range spacing.
5. Manual spacing is the minimum spacing at which drivers feel able to control their vehicles.

The actual values of these quantities for any time and place may be inserted by the system controllers or may be derived from automatic weather recorders, surface friction monitors, and so on.

Hazards

An objective of design is to prevent accidents. An accident, however, has many contributing factors, some of which are random and many of which are outside the designer's control. Therefore, hazards are defined as those factors that, if avoided, will mean that accidents cannot occur. Here, a distinction is made between catastrophic hazards, which are to be avoided in all cases, and critical hazards, which are to be avoided when possible but whose consequences are less serious.

Catastrophic hazards are those conditions that necessarily precede collisions at high delta-V as follows:

- *Hazard A.* A platoon (or single controlled vehicle) is separated from one ahead of it, or from a massive stationary object in its path, by less than platoon spacing.
- *Hazard B.* A vehicle, not under system control, is an unmeasured and unknown distance in front of a platoon or a single controlled vehicle.
- *Hazard C.* A vehicle is released to manual control before the driver has given a positive indication of acceptance.
- *Hazard D.* A vehicle is released to manual control at less than manual spacing from the vehicle ahead of it or at such a relative speed that a spacing less than manual spacing will be realized within, for example, 2 sec.

The following conditions are critical hazards—those that can precede low-delta-V collisions (it is uncertain whether Hazard E should be categorized as catastrophic or critical):

- *Hazard E.* A vehicle under automatic control in the transition lane is less than manual spacing behind a vehicle not under automatic control.
- *Hazard F.* A vehicle within a platoon suffers an electrical or mechanical failure that causes deceleration greater than full platoon braking, does not respond to controls (longitudinally or laterally), or fails to communicate.
- *Hazard G.* A vehicle goes too fast when joining a platoon.

Other hazards exist, such as illegal equipment that can pass false messages to the system controls, interference with the control computers, explosives, heavy weights dropped from bridges, and other deliberate acts. Design features may be desirable to circumvent such activity, but, because they are unlikely to interact with the design of the system as a whole, they are not relevant here.

With these exceptions, these hazards may be sufficient for analysis. That is, if these conditions are avoided, there will be no injury accidents resulting from the actions of vehicles in the system. The system can have no impact on the actions of vehicles it does not control, and these vehicles may be involved in accidents.

Design Consequences

Because the automated lane has a capacity more than twice that of a normal lane, it must be possible to enter it at the side—it cannot achieve capacity if fed from one end only. There is thus no need for a special entrance at the upstream end, which simplifies upstream extension of the system. At the downstream end, however, provision must be made for those drivers who fail to take control from the system (Hazard C). The system must bring them to rest in a specially designated length of the freeway, which is called the dormitory. Vehicles in the dormitory are packed together closely and are moved up as drivers take control of their vehicles and drive off.

In normal operation, at the ends as well as at any intermediate point, entrance and exit will be from a side lane, which is called the transition lane, or TL. Similarly, the automated lane is called the AL.

To avoid Hazard A, no stationary object must be allowed to enter the AL. For example, this situation would happen if an accident occurred on the manually operated lanes and wreckage was pushed onto the AL. Therefore, there must be a barrier between the AL and the TL. Access must be provided either at a limited number of gates or along a limited length of the TL that is shielded from the rest of the freeway by a fence. The latter, however, does not avoid Hazard A; if there were a low-speed collision, the debris might extend from the AL or TL to the other and act as a stationary obstacle. Therefore, only the former case should be considered, as shown in Figure 1. (The arrangement does not avoid Hazard A completely because the gates might admit wreckage from an accident at exactly the wrong place.)

A vehicle will have to make its way across several lanes of traffic before joining a platoon on the AL, and it must do so again when it leaves the freeway. The entrances and exits therefore do not need to be associated with actual on- and off-ramps and need not be confined in length to the length of any additional lane provided at the on-ramp. Indeed, it must not be possible to run out of room on the TL before joining the AL, as it is with physical on- and off-ramps. If none of the gates making up the logical on-ramp (LONR) can admit a vehicle, the driver can revert to manual control and approach the next gate. If a hazard would otherwise arise, a vehicle can also be refused exit at the gates of the logical off-ramp (LOFR). The terms LONR and LOFR encompass not only the physical gates but the total control subsystems associated with joining and quitting the AL.

As a vehicle (controlled or not) enters the gate in the barrier, there is a danger that it will strike the barrier and come to rest partly on the AL. This situation would induce Hazard A, unless the post is instrumented to warn the control system to reduce speeds and unless a vehicle only enters to join onto the rear of a platoon that has just passed. This safeguard would imply that vehicles only enter the AL after they are

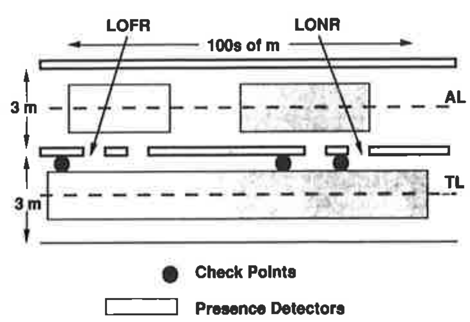


FIGURE 1 Layout of AL and TL.

controlled by the system. It means that Hazard B is avoided, and, by making it possible to test vehicles on the TL before they join the AL, also avoids Hazard F on the AL, where its effects are the most serious.

Thus, the TL must also be equipped to control vehicles whose drivers have indicated a wish to join, at least near the gates. Avoidance of Hazards C and D implies that the TL is also equipped to control a vehicle after it has left the AL. (This equipment includes the LONRs and LOFRs.) Further, with admission only at the rear of a platoon on the AL, LONR capacity will clearly be limited. It seems likely, therefore, that it will be desirable to preform partial platoons on the TL. Hazard avoidance permits this action, and it will be assumed in the following paragraphs.

The previous discussion may have implied that gates are used for entrance or exit, but not both. This restriction may not be necessary, except that the last gate must be an on-only gate. Then, a vehicle that has left the system but whose driver has not resumed control can be readmitted to avoid Hazard C.

The TL will clearly be used by manually controlled vehicles that wish to join or have just left the system. There is therefore no way of excluding vehicles that do not wish to join or are not equipped to do so, and it does not seem right to try. It follows that, to avoid Hazard B, there must be a length of TL, controlled by LONR and LOFR, in which the position of vehicles can be determined by presence detectors. The length will stretch some distance, perhaps 1000 m, upstream of the first gate.

This feature is necessary not only to avoid Hazards B, D, E, and G on the TL but to provide data by which the strategic control system can function. (Strategic control optimizes LONR and LOFR capacity, in part by advising which vehicles should preform platoons, but it is not safety critical.) The presence detectors also make it possible to identify which vehicle is seeking entry and to carry out tests on it (e.g., to ensure that its dimensions correspond to those on its license so that it has no external load).

Requirement Specification of Roadside Configuration

A general specification has been developed for a system that may meet the required criteria of hazard avoidance. The specification consists of the following (see Figure 1):

- An AL will be separated from the rest of the freeway by a barrier. The AL will have a lateral-control reference along its length and will be equipped with a means of communicating with the vehicles on it. There will be an appropriate communication protocol. At its downstream end, there will be a dormitory for storage of vehicles that have failed to resume manual control.
- A series of local control systems, made up of a LONR and a LOFR, will be associated with gates in the barrier. There will also be a length of instrumented TL adjacent to the AL, controlled by the LONR and LOFR. It will stretch upstream from a point a little downstream of the last gate (an on-only gate), to a point perhaps 1000 m upstream of the first gate. The TL will also bear a lateral-control reference and communication equipment. Vehicles entering the AL, pos-

sibly preformed into partial platoons, will join on-AL platoons only at the rear.

- The TL, and that part of the AL immediately adjacent to the gates, will be equipped with presence detectors and a control system that can track identified vehicles through the LONR and LOFR.
- There will be a control system for the AL. Although not necessary, it would seem convenient to associate the length of AL between two access points with the relevant LOFR and LONR to form a block. If so, there will need to be communication between adjacent blocks.
- A centralized strategic control system will propose how speeds will be regulated, through which gates vehicles will enter or leave the AL, and possibly more. It will not be safety-critical because its advice will be mediated by the local block controllers, who will only pass on its suggestions if it will not create a hazard to do so. Any human intervention that may become desirable in normal operation will be through this system. For example, one way in which the local controller may operate is for the vehicle control system to recognize three speeds as relevant to its own speed. One would be the speed of the vehicle ahead, which in a platoon will not be exceeded for very long by very much. The second would be a target speed, set by strategic control, to ensure that, for example, a closing member of a new platoon will come within sensor-range spacing or that the front vehicle of a platoon will reach an on-gate just far enough ahead of a joining vehicle that it can join onto the back. The third would be a maximum speed, set by the local controls to ensure hazard avoidance. It will override the target if demanded by hazard avoidance.
- A body of law regulating permissible behavior on the AL and TL and specifying the equipment that must, may, or must not be present on vehicles will also be part of the system.

Other features may well be present; indeed, those just described are necessary but not sufficient for hazard-free economic operation. From here, it seems that the choices broaden. The specifications presented assume that there will be centralized strategic control (and perhaps self-monitoring), but that safety-critical controls will be local. The actual location of controllers has not been shown to be safety-critical, and may not be, but this arrangement is common on many systems with safety-critical elements.

Uniqueness of Configuration

Beginning with the concept of a single automated lane bearing platoons, and accepting the constraint of hazard avoidance, the configuration of the system is almost completely determined. Therefore it must be generally agreed that safety does indeed demand that the conditions defined as catastrophic hazards be avoided. Simple as they seem, they greatly constrain design choices.

The only other alternative is one in which Hazard B is avoided on the AL, not by insisting that vehicles be under system control before admission but by including presence detectors throughout the whole length of the AL. Such an arrangement would presumably not allow unequipped vehicles to enter the AL legally.

It may be possible to design a system as free of hazards as that discussed here. Avoidance of Hazard E will be difficult and seems likely to make it necessary for control to be passed to the system just before entry; in some reduced form, TL, LOFR, and LONR would continue to be relevant concepts. It seems probable, however, that much of the potential capacity advantage would be lost with such a system. Certainly no reduction in lane width inside the AL barrier would be possible.

However, this possibility is not conclusive and needs further examination. It again seems likely that the constraints imposed by hazard avoidance will be strict.

SYSTEM SPECIFICATION

Operational Modes

It will be a principle of design of any system that, should accidents happen or faults develop, the system will undergo graceful degradation. In other words, the system will pass through a series of modes in which it still operates, though less efficiently, but maintains its hazard-avoidance function. If an accident is severe or the failures are serious enough, vehicles must be brought safely to rest.

A series of operational modes can be envisaged in which any block may operate. The block is the relevant unit. Nothing smaller is possible because all vehicles on the AL in one block must remain there until they reach a gate, and it is clearly not always necessary to take exceptional action in one block because of difficulties in an adjacent one.

In designing a system for safety analysis, seven possible modes of operation have been chosen. This amount seemed initially to be unnecessarily complicated, but no argument has been found that any one of the modes is not required, if any intermediate mode between normal operation and full arrest is to be admitted.

The modes in which traffic continues to move, but at reduced speed, are ones in which all platoons and controlled vehicles on the AL move at sensor-range speed. This speed makes it possible for each platoon to stop safely if there is an obstruction.

The capacity of these modes will depend strongly on the sensor range. As Shladover (*3*) has shown, the capacity of an AL operating with platoon spacings defined as they are here is a function of speed, which has its maximum at 30 to 50 mph. The system controllers will usually set a maximum speed above this limit; however, if the detector range is sufficient, platoons moving at sensor-range speed will give a capacity not less than that at normal speed. Certainly, the loss of capacity may not be unacceptable.

The seven modes are as follows:

1. *Normal.* There are no problems requiring reduced speed or special observation.
2. *Natural.* The system is completely shut down. Vehicles travel on all lanes without any form of system control. This mode needs no hazard analysis.
3. *Sensor-Range–Continue.* Platoons on the AL are constrained to sensor-range speed but, at the end of the block, progress to the next one, which may be in normal mode.

4. *Sensor-Range–Exit.* Platoons on the AL are constrained to sensor-range speed and must leave the AL at the next LOFR.

5. *Crash-Stop.* This mode is used when the presence detectors at the gates indicate a stationary object on the AL or when a signal from the barrier indicates that it has been breached. All platoons brake to rest at full platoon braking. However, if one platoon is less than platoon spacing from, but not at, the breach, each vehicle goes to maximum braking.

6. *Stop.* This mode can only be activated by the system controllers. It is used, for example, when the AL contains paramedics on foot or when debris must be cleared. A number of special functions are possible: a vehicle identifying itself as an emergency vehicle can enter the downstream off-gate and be guided backward down the AL, or a vehicle equipped with special markers (provided by the highway patrol) can be guided backward to the last on-gate of the previous block.

7. *Resume:* Again, this mode is activated only by the controllers. Its effect is to successively accelerate platoons moving in a block on one of the sensor-range modes or the stop mode back to normal-mode conditions, and then to allow the platoon to reenter normal mode.

There are a number of logical links between modes. For example, if a section goes to crash-stop mode, the upstream one must go to sensor-range–exit mode.

Further Development

The seven modes of operation introduce considerable complexity, and a great many choices, into a specification. As operational detail accumulates, the complexities multiply. A system has currently been specified, and fault tree analysis to prove the absence of unforeseen hazards has begun. In the operational area no parallel to the vast simplification that is possible by applying hazards analysis to the physical configuration has been found. When the analyses are finished, they will be reported.

CONCLUSIONS

This preliminary work has revealed a number of results of general importance about AVCS-2 and AVCS-3 systems, with the conclusion that the hazards listed must indeed be avoided.

The possible physical configurations of the system are extremely limited. The hazards do not constrain the form of the system completely, but the number of possible configurations is small and probably includes some that cannot achieve the required capacities. The use of the permitted configurations is necessary but not sufficient for hazard avoidance. There are also operational limitations.

The modular construction of the system and the concept of a strategic controller that is not safety critical, but whose commands are mediated through a local safety-critical controller, seem basic. (These features are also true for many other systems, such as area traffic control.)

The concepts of AL, TL, LONR, and LOFR certainly have wider application and may well be components of all hazard-free systems.

Operations will be of considerable complexity. Command sets and instruction sets are likely to differ among most cells of a matrix of vehicle modes versus modes containing between 50 and 100 cells. It does not seem possible to propose a simple, hazard-free system.

ACKNOWLEDGMENTS

The research reported here is part of the Program on Advanced Technology for the Highway, prepared under the sponsorship of the State of California Business Transportation and Housing Agency, Department of Transportation, and the U.S. Department of Transportation, FHWA and NHTSA. The author would like to thank his colleagues at the University of California (Berkeley), notably Anna Bozzini, Susan Graham, Robert Parsons, and Steven Shladover, for helpful discussion and other assistance.

REFERENCES

1. *Final Report of the Working Report on Operational Benefits.* Mobility 2000, Dallas, Tex., 1990.
2. T. A. Kletz. *HAZOP and HAZAN Notes on the Assessment and Identification of Hazards*, 2nd ed. Institution of Chemical Engineering, Rugby, England, 1986.
3. S. E. Shladover. *Operation of Automated Guideway Transit Vehicles in Dynamically Reconfigured Platoons.* Report UMTA-MA-06-0085-79-1, 2, and 3. UMTA, U.S. Department of Transportation, 1979.
4. N. Leveson. Software Safety: Why, What and How. *Computing Surveys*, Vol. 10, 1988, pp. 125–163.
5. *Procedures for Safety-Critical Software.* Report D-0178A. Radio-Technical Commission for Aeronautics, Washington, D.C., 1986.
6. *Software for Computers in the Application of Industrial Safety-Related Systems.* Report IEC 65A SEC94. International Electrotechnical Commission, Geneva, Switzerland, 1990.

Publication of this paper sponsored by Committee on Vehicle User Characteristics.

Engineering Design Concept for Intelligent Vehicle-Highway System Safety

WEI-BIN ZHANG

The development of an intelligent vehicle-highway system (IVHS) will require methodologies for predicting and controlling system safety. An introductory examination of safety-related issues for IVHS design is presented. Safety problems that may be encountered with the application of IVHS technologies are assessed, and methodologies for safety analysis, system specification, and hazard reduction are examined.

The goals of the intelligent vehicle-highway system (IVHS) and IVHS devices are to improve highway capacity and safety. These two goals are closely interrelated, because an unsafe system or devices would not only cause significant safety problems (liability, societal, or politically oriented) but also dramatically reduce the capacity of the system. Therefore, there is a need for safety studies on the overall suggested IVHS to determine (a) the potential level of safety of the IVHS or IVHS devices and (b) how to accomplish this level of performance.

As explained by Hitchcock in a paper in this Record, studies have been conducted to investigate the safety benefits of the IVHS and IVHS devices by analyzing the frequency and consequences of accidents on the existing highway system and by comparing these incidents with the predictions applied to the IVHS and IVHS devices. The aim of these studies, which support IVHS goals, is to estimate the extent to which the suitably applied advanced technologies will eliminate or reduce human errors in the decision-making and maneuvering process, thereby reducing the occurrence of accidents. These studies are based on assumptions that the system or devices would satisfy predefined safety requirements. IVHS policy makers have described these safety requirements as the prevention of new safety hazards (*1*). A hazard is defined as a condition in which an accident may occur and no action by the control system can prevent it.

To reduce current safety hazards and to eliminate possible new ones, IVHS researchers must conduct an intensive hazard analysis, in which all possible hazards are identified and evaluated. Hazard analysis is the heart of the system safety efforts and can be conducted at multiple levels (or stages), as follows:

- At the system analysis level, major safety hazards are addressed by using preliminary hazard analysis (PHA) and fault tree analysis (FTA). Through these analyses the system configuration can be assessed, and the requirement specifications of the system can be initially defined. The hazards identified are subject to further evaluation and perhaps control in the follow-up studies.
- At the subsystem design level, the preliminary design of components that will perform a specified function of the system is conducted. Detail system hazards are investigated. The cause-consequence sequence of failures and hazard properties is analyzed.

The analyses at the different levels interact in a close-loop form. After a specific safety hazard is identified, hazard reduction analysis must be conducted, including an investigation of the countermeasures for eliminating (or reducing) the hazard in the subsystem design level and improvements (or modifications) to the system configuration and requirement specification in the system analysis level.

The objective of the following paragraphs is to address issues that should be considered in the system safety design. The cause-consequence relationship of system failures is investigated, and the various methodologies for reducing system failures and providing fail-safe features are assessed.

IVHS

An IVHS can be defined as a system that involves the application of sensing, information, communication, and control technologies to observe, guide, and control the movement of vehicles in a traffic system (or to assist in the performance of those functions) (*2*) on existing roadway infrastructures. According to the subfunctions of the system, an IVHS can be grouped into the following four functional subsystems:

1. Advanced traffic management system,
2. Advanced traveler information system,
3. Advanced vehicle control system, and
4. Commercial vehicle operations system.

Structurally, an IVHS possesses component subsystems. Disregarding specific functions, the component subsystems may consist of the following:

- A vehicle system, which may contain communication components, information acquisition or sensing components, information display or warning component, control compo-

Program on Advanced Technology for the Highway, Institute of Transportation Studies, University of California (Berkeley), 1301 South 46th Street, Building 452, Richmond, Calif. 94804.

nent, actuating component, existing vehicle mechanisms, and human-machine interface;

- A roadway or roadside system, which may contain information components, communication components, vehicle presence detection components, and local traffic management components; and
- A central control system, which may contain communication components, network traffic management components, and human-machine interface.

The component subsystems usually interact. Each component subsystem may serve as several functional subsystems, and a group of component subsystems can constitute a functional subsystem. By either classification (functional or constitutional), and IVHS would be hybrid in the sense that the entity contains more than one functional or component subsystem. Figure 1 shows the IVHS component subsystems and their interactions.

There will also be IVHS devices that represent one or more entire or partial component subsystems and perform a single function of an IVHS functional subsystem. IVHS devices will be applied individually before a fully automated IVHS or IVHS subsystem is developed.

DEFINITIONS

In defining IVHS safety, the terms *hazard* and *risk* are used. A hazard in the highway traffic system implies a potential for introducing accidents, which may take the form of a vehicle collision or other undesired events resulting in damage to life or property. Risk is associated with the likelihood or possibility of harm. It is related to the probability that the frequency, intensity, and duration of the stimulus will be sufficient to transfer the hazard from a potential state to a loss (*3*).

According to a generic interpretation, safety can mean hazard free or no risk, as come dictionaries define the term. However, few systems can actually perform at a hazard-free level. Technically, safety can be described using an inverse function of risk (*4*). Hence, a low-risk system will have high safety. In many safety analyses, the degree of safety is measured by using statistics of accidents. Indeed, hazard should be a main factor measured in evaluating safety because it creates the possibility for an accident.

Safety of an IVHS primarily relies on preventing collisions between vehicles because people are the direct users, and because vehicles are actuating components in the system. A vehicle that possesses inherent failures may represent a hazard. It is possible to envisage an accident in which a vehicle collides with a neighboring vehicle or in which an unsafe condition exists that endangers the driver and passengers or others in the vicinity of the vehicle. The unsafe condition refers to the possibility of loss of life or property damage, or both. Therefore, safety of an IVHS or IVHS devices involves ensuring that the vehicles will operate within the designed environments without resulting in an unacceptable risk. Unacceptable risk means that both the frequency and magnitude of the hazards are not publicly (societally or economically) acceptable. There is an arbitrary division between acceptable and unacceptable known as the safety criterion (sometimes defined as safety factor or safety margin). The establishment of a safety criterion for engineering safety has been intensively assessed in risk analysis techniques. The safety criterion can be defined for an IVHS by fatality or property damage and by the influence of the accident on the efficiency of the overall highway system.

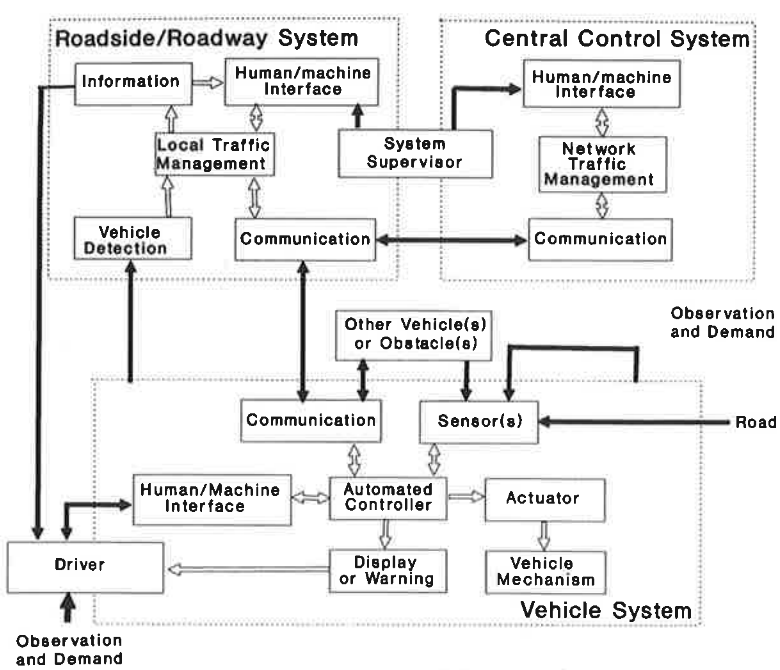


FIGURE 1 IVHS component subsystems.

SAFETY FOR TODAY'S DRIVER-VEHICLE-ROAD SYSTEM

Because it is likely that the IVHS would operate on the existing road infrastructure and in a similar physical environment, it is useful to review the safety problems of the existing highway system.

Drivers, vehicles, and the road have already constituted a system called the driver-vehicle-road (DVR) system. However, whether or not the DVR system is a controllable system is debatable. In the existing DVR system, the driver is the key control component (or subsystem), and the remainder of the system—the vehicle and the road—are, in their present forms, relatively simple components. With regard to the traffic management of the DVR system, the system is sometimes considered as being under control. However, the control functions are effective only at limited locations and only if the drivers obey the traffic regulations and management instructions.

The driver possesses multifunctional sensing and adaptive capabilities; however, these abilities can be restricted by some environmental and human factors. Human sensing abilities are affected by weather or road conditions. In addition, such human factors as alertness, fatigue, and motivation can significantly influence the sensing and decision-making process in vehicle control and may cause operation errors. A driver may be willing to disregard traffic regulations or instructions and may operate the vehicle accordingly. Statistics indicate that driver errors contribute to over 75 percent of all accidents (*5*). Hence, the driver is not considered a controllable operator. Further, it can be concluded that the current DVR system cannot be regarded as a system that is under control in an engineering sense (*6*).

There have been many efforts made toward enhancing safety features for the DVR system, including efforts to design safer cars and improvements to roads and traffic management systems. Clearly, each of these efforts can have, at most, a finite effect, and most of the benefits have already been enjoyed. Improved vehicle and road designs in particular have had positive effects on improving driver performance and reducing human errors; however, these improvements can only lead to limited enhancements against human factor contributions. More advanced traffic management systems may also enhance safety to a limited extent. Hence, substantial efforts have been conducted to provide a means of protection against accidents. The technologies developed for pursuit of this protection are called life-saving (L–S) technologies and largely consist of crush or impact protection designs (*7*). As a result of applying L–S technologies, the structure of the vehicle has been designed to absorb the greatest portion of the energy from most front or rear collisions. Energy-absorbing materials have also been applied for side-impact or roof-crush protection. Safety devices have been developed, including the safety seat belt, supplemental inflatable restraints or air bags, and a self-aligning steering wheel. All of these protective designs have been effective to various extents.

IVHS SAFETY PROBLEMS

IVHS technologies are expected to enhance the safety of the traffic system and to improve the overall system capacity in a number of ways (*8*). By definition, this system should be able to eliminate or reduce human operating error by employing automatic vehicle control systems or hazard warning devices. However, because the IVHS or IVHS devices will be applied on the existing infrastructure, the system or devices must be able to withstand the physical environments that the existing DVR system encounters. Furthermore, the system will also have to confront the operational environments that are developed through the application of the IVHS or IVHS devices, as well as additional influence by human factors.

Safety hazards are considered to be dominated by system failures. Failures may create errors in the operation, and hence, a safety hazard may be introduced. A system failure can be generated by failed components, physical disturbances from the environment, operating errors, errors in the design, or production variations. These sources of system failures should all be considered in the system safety analysis and design. The factors that are considered as stimulus of system failures include component failures, physical disturbances, new operating environments, design errors, and human interference.

Component Failures

As one of the common sources of system failures, component failures may become a dominant factor influencing the safety of the IVHS. Because of the introduction of new components, the system failure rate could be higher than that in the existing DVR system. The electronic, mechanical, and hydraulic components that are introduced in the IVHS must suffer a wide range of environmental conditions. Many of these environmental factors can contribute to component failures, which, in turn, may cause degradation of the functions of the system. The environmental factors may include temperature, humidity, rain, shock, vibration, and so on.

Physical Disturbances

Environmental factors, such as wind (which can create significant external force) and snow (which changes the vehicle-road interaction), will also generate difficulties in the operation of the system.

New Operating Environments

An IVHS or a system with IVHS devices may function in different operational environments than the existing DVR system. For instance, the new system may be designed in such a way to allow the vehicles to operate closer together both laterally and longitudinally, thus improving the capacity of the system. However, if an error that is mainly caused by system failure occurs, and if this error is larger than the designed system tolerance, a hazard will occur. Because the system tolerance of an IVHS would be different from that of the existing DVR system, the specification of hazards suitable for various forms will be different. A similar example can be demonstrated in a system equipped with IVHS devices. If a vehicle has been equipped with a collision warning device, for instance, it is likely that a driver's driving style will change after the driver becomes familiar with the device. By relying on the device, the driver then becomes more aggressive or pays less attention to upcoming objects. If the warning range of the device is set for effective usage on the highway, and if the system then fails to alert the driver in a situation in which a warning should be given, the chance of collision will be greater in an equipped vehicle than in an unequipped vehicle when both are under the same conditions.

Design Errors

Preventing specification or design errors is an important issue in the system design. Many examples have demonstrated, unfortunately, that safety hazards were created by specification or design errors. Design errors often cause safety hazards that are excited by multiple failures or simultaneous failures.

Human Interference

Humans can affect safety in many ways often difficult to foresee: in system production, maintenance, handling, and operation. For example, people sometimes use controls in wrong sequences, do not follow operating instructions, or do not maintain their equipment properly. Therefore, humans may create many difficulties for the system designer and may cause safety problems in the overall system operation.

These factors are fundamentally achieved by combining the stimulus of system failures with enhancement of system hazards. Because these factors exist, it is necessary to specify and design the IVHS or IVHS devices to withstand the environments or to ensure that no (or minimal) undesired events will occur as the result of the interference of the environments.

HAZARD ANALYSIS

Hazard analysis is crucial in designing the IVHS or IVHS devices, because the system configuration and functions can be affected by system hazards. For example, assume a vehicle moving in a platoon undergoes a steering system failure, and the failed steering system turns the vehicle 90 degrees. This situation represents a critical or catastrophic hazard, and the hazard will no doubt lead to an accident. According to this assumption, a system configuration should be designed either to avoid the hazard, which may lead to an accident, or to reduce the collision impact. However, there must be a hazard analysis to investigate under what conditions this particular hazard can occur in the real world. Moveover, the hazard analysis (FTA, in this case) must be used to determine what type of failure would initiate the hazard, whether the failure that causes the hazard would be detectable, and if the hazard would be preventable.

Hazard analysis should be performed in the early stages of safety studies and at the system level. Because the IVHS or IVHS devices contain many unknown factors, it is necessary to make some technical assumptions by defining the system configuration, the environments in which the system works, and so on. The assumptions should also include the component subsystems that will function in the suggested IVHS. Configuring the system is considered a primary step for conducting hazard analysis. At this level, PHA and FTA are usually applied.

When studies on safety hazards are conducted at the component subsystem level, a standard analysis method known as failure mode effects and criticality analysis (FMECA) has to be applied in evaluating the design, method of operation, and environment in which the system works. The cause-consequence relationships leading to system failures must also be identified at this point. Once the component subsystems are defined, major assemblies, subassemblies, or even parts of each component and their functions are then identified. Detailed studies on failure mechanisms of each component (that is, failures, failure modes, and their consequences) then can be examined. FMECA must be followed by FTA to develop the relationship between the occurrence and the sequence of hazard events and finally to evaluate possible failures. FMECA is an essential step in the understanding of the system, without it, FTA cannot be performed.

Systematic safety analysis should be conducted in a closed-loop form. The steps and analytical methods are shown in Figure 2.

ELEMENTS IN SYSTEM SPECIFICATION FOR IVHS SAFETY

Conventional system design usually follows a procedure that includes defining the functional requirements of the system, establishing system specifications, and then designing the system that will accomplish the specified functions. In the system specification the engineers are confronted with an optimization in which capacity and safety are both required. The following elements should be taken into account in the system specification.

Safety Criteria

Safety hazards can be classified as catastrophic, critical, marginal, or negligible (or other classes, if desired), depending on the severity of their consequences. In the system specification, a delineation or safety criterion should be given to define which hazard levels are unacceptable. The safety criterion should take the form of a marginal line on a frequency-consequence diagram, which stipulates the acceptability or unacceptability of the type of hazards on the basis of their consequence and the probability of occurrence.

The establishment of safety criteria must be based on risk assessment at the system analysis level. It requires a systematic consideration of system safety and efficiency, as well as societal acceptability of hazards. A variety of safety criteria for defining the level of acceptability of risks have been proposed for several different applications (*9*). For IVHS, the safety criteria should give, in a quantitative form (e.g., the number of fatalities, severity of injury, or property damages), a description of the types of hazard that are acceptable even if they lead to accidents and the indication of the hazards that are unacceptable. The safety criteria have to be obtained by

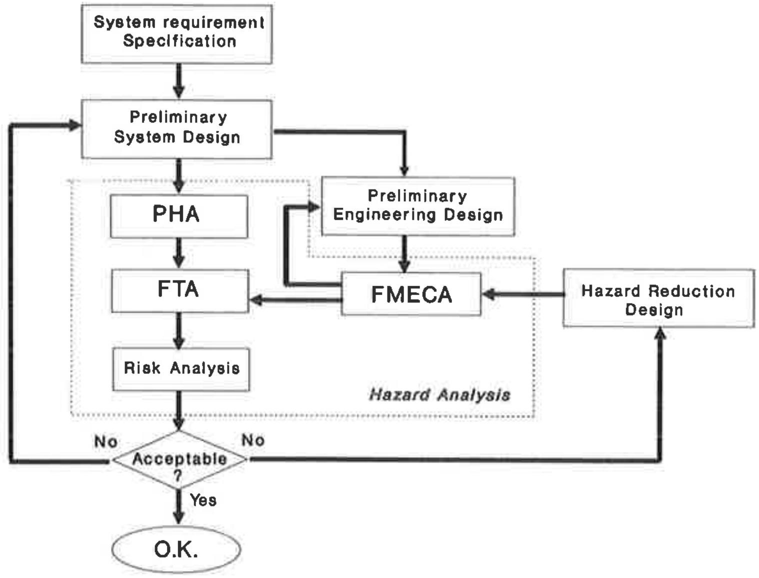


FIGURE 2 Systematic safety analysis.

risk analysis (e.g., the probability risk assessment) in which the impact of all possible accidents within the system is assessed by determining both the likelihood and the consequences of the accidents. Through the analysis, a frequency-consequence relationship can be developed that will first be used in assessing the acceptability of a system and will then serve as a criterion for the system specifications.

Safety States

A system can operate either in normal mode or in failure mode. In the normal operating mode, the system functions as intended in a given condition or situation. (The intended tasks do not include those being wrongly specified, because specification errors have been included in the system failures.) It seems reasonable to suggest that a system is under safe condition if it is in a normal operating mode, because an unsafe condition arises only when a safety-critical component fails to control the vehicle as its design intended. Therefore, in the safety design of a system, attention is usually given to the failure modes of the system.

A failure of a safety-critical component may or may not lead a system into hazards. There would be, according to the predefined safety criterion, only two kinds of states to which the system will be directed as the result of failures:

1. *Fail-Safe (F–S) State.* A state that ensures that no hazard exceeding the safety criterion occurs as a result of system failure; and
2. *Fail-Hazard (F–H) State.* A state in which a hazard exceeding the safety criterion occurs as a result of system failure.

It is the goal of the IVHS researchers to design the system or devices so that it is possible to transfer to an F–S state when hazardous failures occur in the system. Therefore, there is a need to predefine the F–S states according to the discovered failures. Because the severity of a hazard resulting from a particular failure is time or environment dependent, or both, there could be more than one F–S state for a safety-critical component. These F–S states may include graceful degradation, degradation and stop, or emergency stop.

For example, a video collision-detection device may suffer a complete failure of the illumination components. This failure may not affect the system's operation when there is sufficient illumination (natural light or any other kind of illumination). The failure may affect, to a certain extent, the system safety when there is not enough illumination. It will endanger the safety of operation when there is no illumination. Therefore, the F–S states for these three conditions can be normal operation, low speed operation (here, the consideration is concentrating on safety but not on the traffic laws), and stopping the vehicle.

Safety-Critical Components

An IVHS will possess a variety of components at different functional levels. Some components are inherently such that their failures can induce hazards; other are not so. Further, the hazards caused by failures of different components may or may not exceed the predefined safety criterion. While investigating the effects of failures within a particular component or their effects on particular functions, the term *safety-critical* is often used. The safety-critical nature is inherent in the definition of system functions and is incorporated with predefined system functions. Safety-critical functions are those that can induce, cause, or allow a hazardous system state. A safety-critical component is a component or device that performs safety-critical functions. The identification of the safety-critical components must be based on the hazard analysis. To enumerate the safety-critical components, it is necessary to investigate all the possible inherent failures of each component and to examine whether these failures can lead the system into an unacceptable hazard. The safety-critical components should be clearly defined at a very low functional level.

In an IVHS, system safety is defined by the safety of the vehicle, as explained previously. Safety-critical components, therefore, are included among the limited set of components or devices that is directly related to the vehicle-maneuvering functions and the information function. Failures in these functions can introduce system hazards.

Hazard Reduction Principle

In the system specification safety should be incorporated in the definition of the system's physical and operational conditions or environments, as well as hazard control strategies.

In defining the environments attention has to be given to allow the vehicle to operate both in normal and in emergency situations, as well as in situations in which accidents are unavoidable. For example, proper longitudinal spacing between vehicles should be chosen (a) to use the highway effectively, (b) to allow the vehicles to respond correctly to the emergency situation, and (c) to reduce the impacts of the collision if accidents occur.

The system functional requirements should detail the identified hazards, the causes of hazards, and their corresponding F–S states. The system specification also provides a principle for hazard reduction, in which various approaches for transferring a system from a hazard to an F–S state are given. The hazard reduction principle usually follows the system safety precedence, that is, to eliminate hazards, to reduce hazards, to provide safety devices, to provide warning devices, or to provide special procedures.

DESIGN FOR SYSTEM SAFETY

Safety design is essential for providing countermeasures to ensure that both the frequency and the magnitude of accidents do not exceed the acceptable level. To conduct the safety studies, the designer must not only understand and take into account the usual design features, such as operability, quality, and efficiency, but must also fully appreciate the range of environmental conditions throughout the life of the IVHS devices, the range of production methods that will be used in manufacture, and, most important, failures that may occur at any time in the system or within the devices under the predefined conditions.

Many techniques have been developed for providing system safety. In particular, the application of various safety or safety-oriented reliability techniques in transportation engineering, such as techniques applied in railway transportation and air transportation, will have considerable value as exemplars. The safety or reliability techniques are not assessed in detail here; however, several safety design methodologies are discussed.

Design for Robustness

An IVHS or IVHS devices will be applied in a wide range of environments. The components of the system or devices will be produced by different manufacturers, and distinct design and quality control rules will be applied. Therefore, the system or devices will be subject to variations, including allowable production lapses, environmental changes, parameter drift in time, and other factors that can affect the output parameter values in service. Thus, the performance specifications of the system or devices must be met over a large range of input tolerances.

The robust design concept combines elements of control theory and statistical design to optimize product design in relation to its ability to tolerate the expected variations in the environment and production variations. Hence, adaptation of components or the overall system against parameter and process changes is provided. With robust design application, production and maintenance can be relative easy and less expensive.

Design for Reliability

The reliability design concept is applied to ensure that the system continues operation under given conditions for its expected life. Through reliability design, a component should possess sufficient ensurance against progressive weakening to withstand fatigue, corrosion, wear, and so on, thus reducing its failure frequency.

By applying appropriate reliability techniques, it would be possible to design a system with a low failure rate according to the strictly defined specifications. However, whether an IVHS or any IVHS devices can be constructed to meet and be maintained at specific engineering requirements is doubtful. It is well known that the failure rate of any system may be substantially higher than expected in the original design according to the so-called burn-in and wear-out phenomena (which correpond to the initial failure rate and aging failure rate in the "bathtub" curve). In reality, it is obviously not economically practical to design and produce a huge number of vehicles that will always be inherently failure free in all environments. In addition, there is always the possibility of having errors in the design. The practical considerations of cost might also limit the extent to which total reliability can be ensured. For these reasons, reliability design will be applied as one of the design methodologies for system safety, but not the primary methodology.

F–S Design

Fail-safe, as the term suggests, should indicate the ability of a system to ensure that it can be handled safely (i.e., to avoid a hazardous condition) should it reach conditions outside specified tolerances.

As a design principle, F–S has sometimes been interpreted as fail-stop; therefore, this approach has been considered to be inefficient. In fact, this interpretation involves distortions. By using the definitions of system states, the F–S property can be redefined as the ability of allowing a system to enter into an F–S state in the event that the system deviates from the normal operating mode. Because the F–S states are defined in accordance with the particular failure modes of the components, a system that is designed on the basis of the F–S principle will lead a system into a corresponding F–S state, including maintaining the system operation or one-level functional degradation (e.g., reduce the vehicle speed). It is true that fail-stop will be applied, because stop is the single, final degradation state of the F–S states; however, it is applied only when other functional degradation states would not be enough to prevent the system from unacceptable hazards.

F–S design can be accomplished by a specially designed protection system that can detect system failures in the early stages and then lead the system into previously defined F–S states. The key issue in F–S design is the assurance that the protection system will be fail-safe. Two types of failures are addressed in designing the F–S protection system: (a) F–H failures (where the protection system fails to guide the system into an F–S state when one or more inadvertent events occur and (b) F–S failure (the protection device forces the system to enter into an F–S state when no inadvertent event exists). It is crucial that the protection system be designed fail-safe, that is, to avoid F–H failures, although F–S failures also need to be reduced.

Other Safety Design Approaches

Other safety approaches should also be considered in the safety design to reduce the severity of impacts if vehicles in the IVHS are involved in accidents. Many L–S technologies that have been employed in the existing DVR system can be applied to protect people from fatality, injury, or property damage.

CONCLUSIONS

Designing for minimal acceptable hazard in an IVHS or IVHS devices means that a complete identification of system hazards will have to be accomplished from the hazard analysis. Then, alternatives for eliminating or controlling the specific hazard will have to be evaluated so that an acceptable control method for hazard reduction can be chosen.

To identify the safety hazards completely, preliminary design of components of the system or devices is required to understand in depth the failure mechanisms of the system. Thereafter, FMECA and FTA can be conducted. Safety criteria should be given in the system specification and design; thus, the safety-critical components and safety states of the IVHS or IVHS devices can be defined. The strategies that will lead the vehicles to safety as the result of system failures should be designed for those components that are defined as safety critical.

Important issues for designing a safe IVHS or IVHS device have been investigated, and some preliminary definitions have been given. Technological concerns in the specification and design of the IVHS or IVHS devices have been discussed. The design methodologies for providing system safety have been assessed to provide a basis for evaluating technologies applied for reliability and safety design. Further work will be reported later.

ACKNOWLEDGMENTS

The research reported here is part of the Program on Advanced Technology for the Highway, prepared under the sponsorship of the State of California Business Transportation and Housing Agency, Department of Transportation. The author also wishes to thank Anna Bozzini, Anthony Hitchcock, and Steven Shladover for their contributions, technical or otherwise, that helped to shape the final form of this paper.

REFERENCES

1. *Final Report of the Working Group on Advanced Vehicle Control System (AVCS).* Mobility 2000, Dallas, Tex., March 1990.
2. S. Shladover. Roadway Automation Technology—Research Needs. Presented at 68th Annual Meeting of the Transportation Research Board, Washington, D.C., 1989.
3. H. Roland. *System Safety Engineering and Management.* Wiley, New York, 1983.
4. P. O'Connor. *Reliability Engineering.* Hemisphere Publishing Co., Washington, D.C., 1988.
5. I. Treat et al. *Tri-Level Study of the Causes of Traffic Accidents—Executive Summary.* Report DOT-HS-805-099. U.S. Department of Transportation, May 1979.
6. J. F. M. Bryant. Is the Motor Vehicle and Its Drive Part of a System? Human Factors in Transportation. *Proc., 10th Annual Conference, Ergonomics Society of Australia and New Zealand.* Adelaide, S. Australia, 1973.
7. R. Stempel. Where Auto Safety Stands Today. In *Media Update on Passive Restraints*, Warren, Mich., July 1989.
8. E. Deakin. *Opportunities and Constraints for Advanced Highway Technologies: A Speculative Analysis.* Research Report UCB-ITS-PRR-89-7. Institute of Transportation Studies, University of California at Berkeley, Richmond, Oct. 1989.
9. J. Thomson. *Engineering Safety Assessment: An Introduction.* Longman Scientific and Technical, 1987.

Publication of this paper sponsored by Committee on Vehicle User Characteristics.

Real-Time Responses to In-Vehicle Intelligent Vehicle-Highway System Technologies: A European Evaluation

NICK AYLAND AND JAMES BRIGHT

A gaming and simulation experiment is being conducted in Europe to assess the behavioral impacts of in-vehicle intelligent vehicle-highway system (IVHS) technologies. The experiment is being carried out as part of the EURONETT project under the European DRIVE initiative. It concentrates on real-time responses to in-vehicle driver information systems within an urban environment. A brief overview of the DRIVE program is provided. The objectives and scope of the EURONETT project are described, and the role of simulation and experimentation within EURONETT is explained. The experimental design and methodology are then presented, and the measures of effectiveness being used are described. Finally, the innovative gaming approach being adopted is summarized, and application of the results is considered.

The world of road transportation is undergoing a fundamental change. New technology is being developed at a great rate, which can help motorists, transportation authorities, and fleet operators make better use of their vehicles and the highway network. These advances are particularly important in view of the rising urban congestion levels being experienced worldwide.

The potential of these new technologies and growing concern over road transportation problems have led to a number of research, development, and demonstration initiatives throughout the world. These initiatives include the U.S. intelligent vehicle-highway system (IVHS) program, European programs such as DRIVE (Dedicated Road Infrastructure for Vehicle Safety in Europe) and PROMETHEUS, and Japanese programs such as RACS and AMTICS. These initiatives embrace a wide range of technologies, such as traffic management and control systems, driver information systems, and automated vehicle control systems.

However, the success of new technologies in improving road transportation will depend on policy makers and industry choosing the most appropriate technology and applying it in a way that will maximize its benefits. Knowledge of the ways in which motorists are likely to respond to different types of technology is required. This crucial issue is being addressed within EURONETT (Evaluating User Responses on New European Transport Technologies), as part of the European DRIVE initiative.

N. Ayland, Castle Rock Consultants, 18 Liberty Street, SW, Leesburg, Va. 22075. J. Bright, Castle Rock Consultants, Heathcoat Building, Highfields Science Park, Nottingham NG7 2QJ, England.

One particular aspect of EURONETT—an experiment to assess real-time behavioral responses to in-vehicle driver information systems—is addressed. This experiment covers responses to on-board navigation aids, dynamic route guidance systems, and the radio data system traffic message channel (RDS-TMC) traffic information broadcasting system. Following an introduction to DRIVE and EURONETT, a description of the innovative experimental methodology being used and the measures of effectiveness being recorded are provided. Application of the experimental results is then considered, and conclusions on the work being carried out in this vital area are presented.

DRIVE

The DRIVE program is a $140 million research program of the Commission of the European Communities (CEC), linking information technology and transportation. DRIVE was formally adopted as a community research program in June 1988, and 61 cooperative projects were subsequently funded under the first program phase.

DRIVE has three primary objectives, as follows (*1*):

1. Improving road safety,
2. Maximizing road transport efficiency, and
3. Contributing to environmental improvements.

DRIVE envisages a common European road transport environment in which drivers are better informed and intelligent vehicles communicate and cooperate with the road infrastructure itself. The program follows a top-down systems approach to the research and overall design of traffic management and safety systems, which represent a significant advance over those now available.

Specifically, DRIVE aims to achieve the following results:

- Identification of the best choice of systems on the basis of economic, social, and technical criteria;
- Development of optimal strategies for their implementation;
- Specification of performance and compatibility standards that will enable industry to develop the necessary equipment and systems;
- Provision of directives and guidelines to which industrial products and intelligent European road transport infrastructures should conform; and

• Design and implementation of pilot schemes to assess the performance of equipment and systems.

EURONETT

EURONETT is a 3-year project that lies at the heart of the DRIVE program (2). Its main objective is to assess and predict the behavioral responses of travelers to policy initiatives on the basis of new technology (or the IVHS). Within this framework, it also aims to produce a EURONETT toolbox comprising various models that can be used in the future by transportation authorities in evaluating the likely effects of alternative measures.

EURONETT feeds into many other projects currently in progress, both within DRIVE and in national or local projects taking place in individual European countries. For these projects, it will provide vital data and tools for assessing effects that assume a new importance with the introduction of IVHS technologies. Equally, some of the products of EURONETT will ultimately be applicable to similar investigations in North America and Japan.

The EURONETT project is being undertaken by a consortium of six partners from four countries, including government user organizations, consultants, and universities with relevant expertise. The project consists of six major tasks, which are structured into eight work packages as shown in Figure 1. The major research investigations are concentrated in Work Package 3, whereas Work Packages 5 and 6 involve development of the EURONETT toolbox.

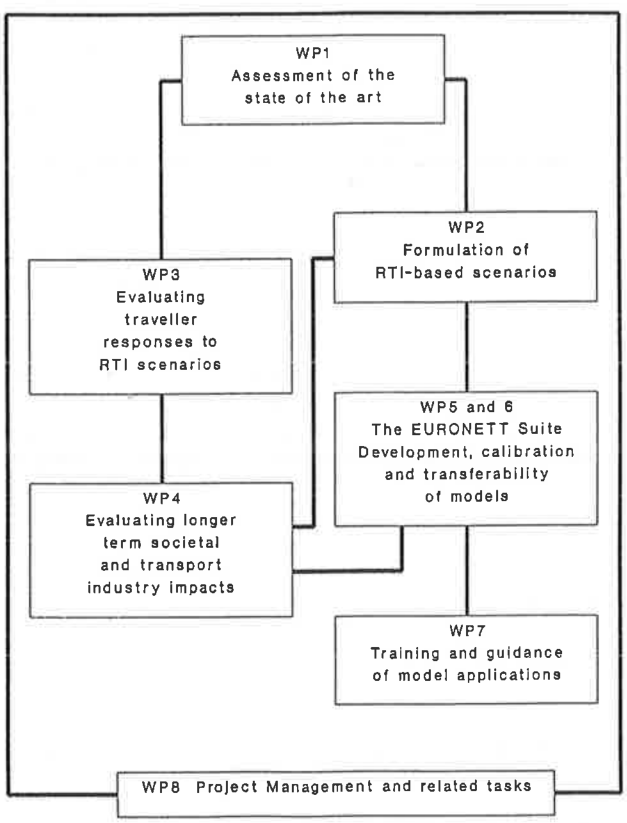


FIGURE 1 EURONETT project structure.

IN-VEHICLE SIMULATION AND GAMING

Within the EURONETT project, much of the research into user responses to new technologies is being conducted using established stated preference techniques. Surveys are in progress using these techniques in two European cities (Athens and Birmingham) to investigate user responses to a range of technologies. These include the following:

- Electronic road pricing systems,
- Improved public transport information systems,
- Advanced integrated traffic control systems, and
- Electronic trip planning services.

These surveys have proved successful in evaluating user reactions; although the technologies themselves may be complex, interactions with system users are relatively straightforward. However, in-vehicle driver information systems often require the driver to make route-choice decisions in real time while undertaking a task (driving) that is itself onerous and demanding. The complexities of this decision-making process are difficult for someone to perceive when removed from the driving environment and presented only with visual stimulus material.

Therefore, to investigate user responses to in-vehicle driver information systems, an alternative approach has been adopted with EURONETT. This approach uses a simulation and gaming technique, which makes the experimental subject act on information provided by these in-vehicle systems in real time. The subject simultaneously performs a task that simulates driving a vehicle through an urban street network to represent the situation that would be encountered in using such systems on the highway.

The main reason for using gaming simulation is to create or re-create decision environments to observe how people assess options and reach decisions in situations that are either familiar or unfamiliar to them. It is the creation of the decision environment that strongly influenced the adopted approach.

The decision environment present in a driver's navigation task was examined, and it was apparent that certain key features of this environment must be included in the simulation, as follows:

- A real-time environment,
- A task of subjective mental workload commensurate with driving,
- A task containing navigational decisions, and
- A task allowing the presentation of driver information conceptually compatible with European road transportation technology initiatives.

After a review of possible ways to create this decision environment, a computer game was selected that fulfills these criteria. However, it was recognized that a validation exercise would need to be included within the experimental framework to ensure that behavior in the simulation trials is similar to real-world driving and navigational behavior. Because of external factors relating to the project, the validation exercise is being undertaken after the main series of simulation trials. It is discussed briefly in a later section.

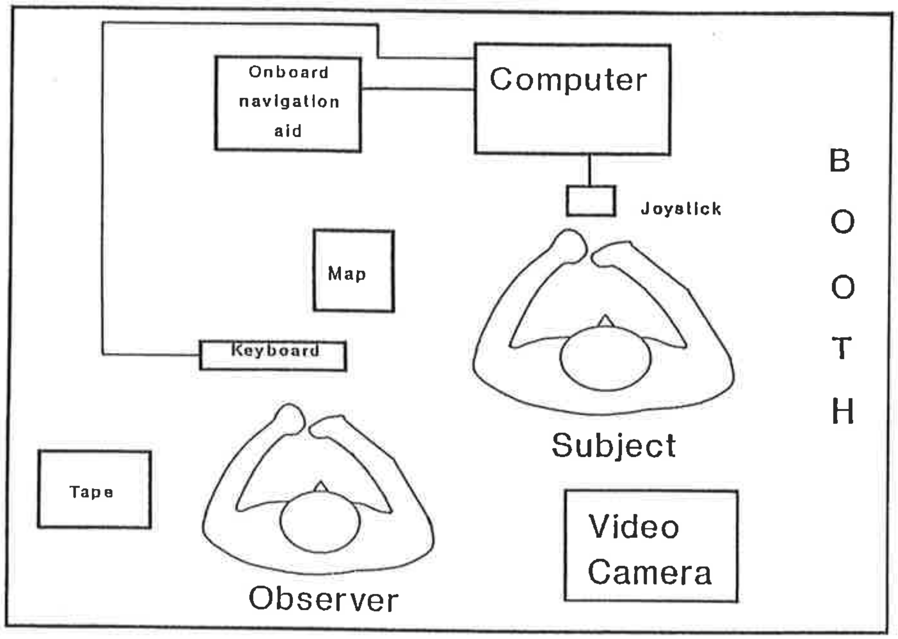


FIGURE 2 Experimental layout.

EXPERIMENTAL TOOLS

The primary objectives of the gaming and simulation work in EURONETT is to get a measure of subjects' route choice decisions in an urban network when provided with different types of in-vehicle driver information systems. The laboratory simulation approach is intended to complement the other approaches being adopted in EURONETT to assess drivers' responses to new IVHS technology.

The experimental design makes use of a commercially available computer game. This game provides a three-dimensional simulation of the San Francisco road network and gives a driver's-eye view of the road ahead and the vehicle controls and gauges. The experimental subject is then able to navigate the vehicle through the network, while being presented with information from different in-vehicle driver information systems.

This computer game was selected as the basis for the simulation and gaming experiment following pilot trials. The software runs on an IBM PC adapted with an enhanced graphics adaptor (EGA) and a joystick interface and card. For the experiments, the software is being used in training mode, which precludes racing or serious accidents.

Control of the subject's car can be overridden from a keyboard so that the car can be slowed by the experimenter. By limiting the driver to first gear, a maximum speed of 40 mph can be achieved. This limit is realistic for the urban environment being used for the experiments. Piloting of the software using young drivers previously unacquainted with the software showed that 10 min is a reasonable training time for subjects to achieve a sufficient level of confidence in controlling the vehicle and avoiding collisions.

The software models the network in three dimensions in real time using EGA color graphics and sound effects. There are various viewing modes. The mode selected for the experiment allows a driver's view through the windshield, giving a good three-dimensional representation of the vehicle gauges and steering wheel and realistic optic flow patterns generated by oncoming roads and passing buildings. The software includes other traffic, such as buses, trucks, ambulances, taxis, motorbikes, and police patrol cars. There are also pedestrians on paths and crossing roads.

The area of the network selected for experimentation starts by the San Francisco Zoo on the southwest side of the city and extends to a gas station one block down from Golden Gate Park. The test network therefore incorporates four distinctive landmarks. The zoo is at the starting point. On the left-hand side, traveling toward Golden Gate Park, is the Pacific Ocean. Next is Golden Gate Park itself, and at the destination there is a gas station. Along the route there are secondary landmarks in the form of differently colored buildings.

During the experiments, the subject sits directly facing the PC screen to the simulation game. Connected to the computer is a joystick, which is fixed in front of the subject. A second screen is located adjacent to the main computer for use in one particular technology scenario. The experimenter is seated out of the subject's line of sight and has a tape recorder and computer keyboard within reach. A video camera is positioned behind and above the subject to get a view over the subject's shoulder so that all key experimental parameters can be recorded. This layout is shown in Figure 2.

EXPERIMENTAL DESIGN

Experimental Procedures

The experiment is essentially a "within-subjects" design. In the first phase being conducted within EURONETT, a relatively small subject sample is being used, with up to 20 experimental participants. The participants have therefore been selected to have similar demographic characteristics, although some initial examination of variation in response by age range is included. In view of the limits on resources available, the intention in this initial experimentation phase is to concentrate on one particular population group: young drivers 17 to 30 years old. Subsequent experimental phases (possibly outside the current EURONETT project) would further investigate variations in response across different demographic groups.

Each experimental subject is given the task of navigating the vehicle to the gas station destination on each of 20 repeat trial runs. At least 12 route-choice decisions must be made on each run. During each run the subject is given different items of information en route to represent one of four technological scenarios (described later). Each subject therefore attempts the journey five times under each scenario, with the order of presentation of the different scenarios being randomized.

Before the trials commence, subjects are given a 15-min period to familiarize themselves with the driving task on a separate part of the network. During this familiarization period, subjects experience congestion and delays and get to see an example of the landmark they are attempting to reach.

Subjects are paid an incentive for participation, with a larger prize offered to ensure sensible participation. Subjects are encouraged to

- Drive safely,
- Drive legally,
- Drive along a sensible route, and
- Drive swiftly.

These criteria are also the ones used to judge the best overall driving performance in awarding the prize.

During the course of each trial, the experimental subject may encounter delays or congestion on the road network. This situation is characterized within the simulation game by an external slowing of the subject's vehicle to 0 to 5 mph for a predetermined duration. Two levels of congestion severity are included within the experimental design, each having a different congestion duration. Congestion patterns across the network are selected randomly by the experimenter from 20 defined patterns before each trial (without the knowledge of the experimental subjects).

At the conclusion of the experiments, subjects are fully debriefed. The experimenter answers any questions the subjects may have. In addition, a short questionnaire is used to elicit subjective impressions of the value of each technology in making navigational decisions.

Technological Scenarios

Four technological scenarios are being investigated within the gaming and simulation strand of EURONETT: one baseline scenario and three scenarios focused on different in-vehicle driver information technologies.

Scenario 1

The baseline scenario represents a normal driving situation, in which a driver has no technological aids to use in navigating around an unfamiliar highway network. Information available to the subject at all times includes a street map of the area, plus the name of the street on which the vehicle is traveling and the cross street being approached. The street name information is provided through a facility on the dash display in the driving simulation game. Information that could normally be gleaned from street signs is given. This minimum level of information is also available to subjects in the other scenarios.

Scenario 2

The second scenario provides the driver with an on-board navigation aid in the form of an in-vehicle electronic map display. A street plan is illustrated on a second small computer screen, and the current position of the subject's car on the network is indicated by a flashing blue mark. This display operates in real time and is constantly available to the driver. The display is located to the side of the driver at a viewing angle that is typically of in-vehicle dash mounting.

Onboard navigation aids (*3*) with electronic map displays are being developed and introduced by a number of manufacturers in Europe and other areas of the world. The additional help they can provide could potentially affect route-choice decisions, justifying their inclusion in the experimental scenarios. Current systems that use some form of locational display include ETAK and the Philips CARIN system.

Scenario 3

The third scenario covers the use of an externally linked route guidance system (*4*). Such systems include electronic route-planning and route-following aids, which have a communications link from in-vehicle guidance equipment to an external system providing real-time network or traffic information. The principal difference between these systems and self-contained electronic map display aids is that they provide actual routing advice to the motorist, which can account for real-time traffic conditions.

In this scenario, a route guidance system is represented that is similar in concept to the AUTOGUIDE and LISB systems being implemented in pilot schemes in London and Berlin, respectively (*5*). These systems use a network of infrared beacons strategically located on the road network, as well as in-vehicle units with an in-built dead-reckoning capability for navigating between beacons. Two-way data communication between the in-vehicle units and the beacons allows information on network conditions to be exchanged between the vehicle and a real-time network condition data base.

In trial runs under this scenario, information is presented to the subject in the form of audio messages at key intersections. These routing advice messages conform to the AUTOGUIDE specification developed for the London pilot scheme. For the trial street network, the messages guide the subject along an optimal or near-optimal route, calculated according to the congestion pattern for that particular trial run. Provided the subject follows the guidance advice, advice is provided at all subsequent key points.

Scenario 4

The fourth scenario provides simulated information from the RDS–TMC. The RDS is a system that enables digitally encoded data to be inaudibly superimposed on the stereo multiplex signal of a conventional FM radio broadcast (*6*). It is currently being widely introduced in Europe and has been chosen as a world standard.

International efforts are taking place to develop standards for the RDS-TMC (*7*). A specialized receiver is required to decode traffic messages, which can then be displayed either as text or as synthesized speech. Such a system is particularly attractive from a European perspective, because digital transmissions can be interpreted in the language of the driver's choice. Studies are also being undertaken in countries outside Europe, including Canada and Hong Kong, to investigate the possibilities of implementing RDS-TMC. In the United States, Chrysler has been carrying out trials of RDS traffic information in the Detroit metropolitan area.

In trial runs under this scenario, a sequence of audio RDS-TMC messages is presented to the subject at the start of each run. These messages conform with standard message formats currently being finalized in Europe and relate to the congestion pattern for that particular trial run. The messages can be repeated on request en route to represent that capability of RDS-TMC receivers.

EXPERIMENTAL MEASURES

There are many measures that could be taken using the experimental design and procedure that have been adopted. However, to keep the analysis within reasonable limits, to maintain direct relevance to the experimental objectives, and to minimize the load placed on the experimenter, only the most important subset was selected.

The measures in this work fall into two categories:

1. Primary measures, and
2. Secondary control measures.

Primary Measures

The primary measures relate directly to route choice decisions made by subjects. These measures were developed by considering the different ways in which each technology aims to improve drivers' navigational decisions. The baseline scenario provides basic information on the street network but no real-time traffic condition information. Scenario 2, with an onboard navigation aid, provides an additional tracking capability but again no real-time information on traffic conditions. Drivers avoiding congestion by using these information sources must therefore do so intuitively.

Scenario 3, involving a route guidance system, provides actual guidance advice on the basis of knowledge of real-time traffic conditions but offers no direct information on the conditions that cause the advice to be given. The effectiveness of the system lies in how often advice from the system is accepted and followed. Conversely, RDS-TMC (Scenario 4) provides information on traffic conditions but no explicit advice on optimum routings. The system works by persuading drivers to avoid congested areas about which messages have been broadcast.

With these considerations in mind, primary experimental measures were included in the experimental design that will provide information on the percentage of drivers who will act on real-time information or advice given (Scenarios 3 and 4) or who will avoid congestion intuitively (Scenarios 1 and 2). This information can then be used as input to route-choice models. These primary measures are as follow:

- *Scenario 1.* Number of times congestion is encountered ÷ number of trials;
- *Scenario 2.* Number of times congestion is encountered ÷ number of trials;
- *Scenario 3.* Number of times routing advice is taken ÷ number of trials; and
- *Scenario 4.* Number of times a subject hits congestion about which there has been an RDS message ÷ number of trials.

A further important primary measure was included within the experimental design to provide a comparison of the quality of route chosen by subjects under each of the technological scenarios. Although route quality is difficult to measure definitively, a measure called route quality index (RQI) was defined for comparison of scenarios. The RQI is based on the difference between a subject's chosen route in each trial and the optimum route for that trial, accounting for the prevailing congestion pattern.

In calculating the RQI and in determining optimum routes, consideration was given to research on drivers' criteria in route selection. This research (*8*) indicates that trip time and trip distance are dominant factors in drivers' route-choice decisions, with trip time being the most important factor. Hence, some balance of these two factors should be used in identifying optimum routes and in judging the quality of routes selected by subjects.

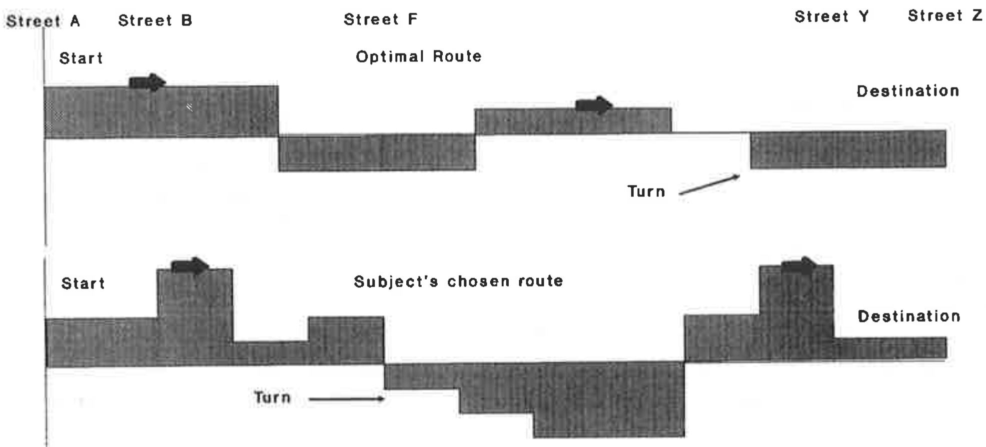
The RQI used in the experiments is therefore a weighted sum of the difference in trip time and the distance between the subject's chosen route and the calculated optimum route, as shown in Figure 3. Route quality is therefore calculated as follows:

$$\mathrm{RQI} = a \times \frac{\text{trip time (chosen} - \text{optimum)}}{\text{trip time (optimum)}} + b \times \frac{\text{trip distance (chosen} - \text{optimum)}}{\text{trip distance (optimum)}}$$

The constants a and b can be varied to reflect the relative importance of each variable in the overall RQI. Using this measure, it follows that the optimum route has an RQI that tends to zero. The value of the RQI becomes progressively more positive as the route progressively deviates from the optimum.

Secondary Measures

The secondary control measures recorded within the experiment are designed to provide feedback on the performance



$$\text{Route Quality Index (RQI)} = a \times \frac{\text{trip time (chosen} - \text{optimal)}}{\text{trip time (optimal)}} + b \times \frac{\text{trip distance (chosen} - \text{optimal)}}{\text{trip distance (optimal)}}$$

FIGURE 3 Measurement of route quality.

of the experimental design. These measures were identified as the most likely to provide explanations of some of the more plausible causes of variance in the final results. They also provide objective measures of subjects' continued serious participation.

The first of these control measures is the time taken by each subject to complete each trial. This time is measured to check for learning effects on the network. If there is a significant decrease in time taken over the trials, then it will be possible to conclude that learning has had a major effect. This measure also helps to indicate whether subjects are treating the experiments as a real-time exercise or are regularly taking time to solve what they perceive as the problem. Finally, this measure taken with any particular congestion pattern provides an index of difficulty for that pattern.

The second control measure is the number of times the map or the onboard navigation aid is scanned. This measure will provide a rough indication of whether subjects really are using the system. Piloting indicated a mismatch between the subjects' verbal reports during debriefing and their actual trial behavior, which is not uncommon in psychological research.

The third control measure is the number of times the RDS messages are requested over the trials. This measure provides a rough indication of whether subjects are using or attemping to use the RDS.

Finally, the number of collisions is recorded. Although most collisions do not affect the progress of the trial run, this measure provides an objective evaluation of the subjects' safe driving. It enables a rough guess as to whether the subjects are (a) competent and (b) taking the experiment seriously.

FUTURE WORK AND APPLICATION OF RESULTS

Once simulation and gaming trials are completed, a full analysis of results will be conducted. These results will provide valuable insight into the behavioral responses of motorists to in-vehicle IVHS technologies, as well as a comparative measure of the value of the each type of technology under test.

On completion of the laboratory-based work, a validation exercise will be undertaken to ensure that the simulation results give a true measure of real-world driver responses. This exercise will use an instrumented vehicle, in which a limited subject sample will be given navigational tasks comparable to those of the simulation experiments. The validation trials will take place on a restricted area of the actual highway network and will encompass each of the technological scenarios.

Within the validation exercise, data will be collected that will allow tests for realism to be conducted. For example, the frequency with which information sources are scanned or requested in the simulations and in the validation trials will be compared. Similarly, glance duration and other measures recorded on video in both sets of trials will be compared.

The results of the simulation experiments will be used to modify and calibrate a route choice model. This model will form part of the EURONETT toolbox, which as a whole will allow the effects of different levels of behavioral response to new technologies to be assessed on the European road transportation system.

The route-choice model will allow a future vehicle fleet to be represented, with different types of technology in various percentages of vehicles. The model can then be used to examine the effects on traffic performance of an incident or congestion occurring on the network. The results of the simulation trials will indicate the rerouting responses for each type of technology within the vehicle fleet, and these responses can then be used to compute overall network effects.

This facility will be valuable for making future policy and planning decisions within the European transportation environment. Application of the model in the United States would also give a powerful tool to transportation agencies and legislators, if U.S. behavioral data could be elicited from a similar experimental program. This is particularly so in view of the IVHS initiatives now gathering momentum in the United States.

SUMMARY

A description has been provided of an experiment currently in progress in Europe to determine behavioral responses to in-vehicle IVHS technologies. The experiment is being undertaken as part of the EURONETT project under the European DRIVE initiative. It uses an innovative PC-based simulation and gaming approach to focus on real-time responses to in-vehicle driver information systems within an urban environment.

The background to DRIVE and EURONETT has been briefly outlined, and the way in which the simulation and gaming work fits in with other EURONETT investigations has been described. Experimental tools and design have been discussed, together with measures being recorded within the investigation. Future work and potential applications have also been outlined.

The technologies being investigated within the experiment include onboard navigation aids, route guidance systems, and the RDS-TMC. These system categories were selected for the following reasons:

- All are in-vehicle driver information systems, which therefore require routing decisions to be made while the normal driving functions are being undertaken; and
- All are almost certain to be in use in the near future in the early stages of a European integrated road transport environment.

The results of the simulation and gaming experiment will be of value in assessing the traffic performance effects of the widespread introduction of these IVHS technologies in congested urban areas. They will permit modification and calibration of route choice models, allowing wide area effects of incidents and congestion with a future IVHS-equipped vehicle fleet to be determined. These models will form powerful tools for transportation agencies in both Europe and the United States allowing IVHS implementation to proceed during the next decade on a fully informed basis.

ACKNOWLEDGMENT

The EURONETT project is being undertaken by a consortium of partners on a cooperative basis. The input of all partners to the project is duly acknowledged.

REFERENCES

1. *R&D in Advanced Road Transport Telematics in Europe—DRIVE 90*. Commission of the European Communities, Brussels, Belgium, March 1990.
2. *Evaluating User Responses on New European Transport Technologies—1989 Annual Report*. EURONETT Consortium.
3. D. J. Jeffrey. Electronic Aids to Vehicle Navigation—An Overview. *IEE Digest*, Vol. 21, London, 1987.
4. N. D. Ayland and P. Davies. Advanced Technology for Relieving Traffic Congestion. *Proc., Institute of Transportation Engineers Conference: Traffic Congestion 1988—Issues and Answers*, Tampa, Fla., 1988.
5. R. Von Tomkewitsch. LISB—Large Scale Test Navigation and Information System Berlin. *Proc., PTRC Seminar on Information Technology in Transport and Tourism*, University of Bath, England, 1987.
6. *Specification of the Radio Data System (RDS) for VHF/FM Sound Broadcasting*. 3244-E. European Broadcasting Union, Brussels, Belgium, March 1984.
7. *Review of Progress to Date on RDS-TMC*. RDS-ALERT Consortium, June 1989.
8. J. Wootton and M. Ness. *The Experience of Developing and Providing Driver Route Information Systems*, IEEE, Washington, D.C., 1989.

The views expressed in this paper are those of the authors and do not necessarily reflect the views of the EURONETT consortium as a whole.

Publication of this paper sponsored by Committee on Vehicle User Characteristics.

Consumer Acceptance of Adaptive Cruise Control and Collision Avoidance Systems

THOMAS TURRENTINE, DANIEL SPERLING, AND DAVID HUNGERFORD

Consumer reactions to automated vehicle control technologies were studied. The motivating hypothesis was that current users of cruise control value the relaxation benefits they gain from its use and would therefore be early adopters of more automated controls. Four focus groups were conducted, two with avid users of cruise control and two with infrequent users. The hypothesis was not borne out: avid users valued cruise control as a driving aid more than as a means to relax and thus had little interest in more advanced automated controls. Less frequent users, in contrast, were more attracted to the automated controls because of the increased safety benefits they could provide in emergencies, although the users expressed concern about reliance on those automations in inappropriate circumstances. It is hypothesized that (a) avid cruise control users are not a special early market; (b) safety is the primary feature, both negatively and positively, in defining the early market; and (c) convenience is not likely to be a primary feature attracting early adopters of automated driving controls.

Saxton (*1*) refers to the transition between manual and automated control as a "fundamental evolutionary gap," suggesting a determined historical process rather than an open, market-driven process. Combining the concept of a linear development path with a market path, Johnston et al. (*2*) suggest that the implementation of intelligent vehicle-highway system (IVHS) freeways will take place as a five-stage process, whereas Chen and Ervin (*3*) envision a nine-step process. In both cases the first stage or stages would be voluntary, market-driven purchases of onboard navigation and route guidance systems. The second set of steps would be voluntary, market-driven purchases of automated acceleration and braking systems, which use radar-like sensing systems to warn drivers about possible collisions and to decelerate the car automatically. Other automatic vehicle control systems (AVCSs) also in development are automated steering systems and proximity sensors to warn of sideswipes and backup accidents (*4*).

Deploying IVHS technologies requires public investment in intelligent highways and private investment in intelligent vehicles (*5*). Public investment will require public approval, whereas private investment will require consumer interest.

The primary purported benefits to individual drivers from the IVHS and their motivation for purchasing IVHS technologies will be reductions in driving stress due to better information on travel conditions from advanced vehicle information systems (ADISs), increased safety from advanced vehicle control systems (AVCSs), and shorter travel times due to congestion reduction (*6*). However, the majority of these benefits will accrue to users only when the IVHS is fully deployed. Therefore, early consumer demand for ADIS and AVCS components is less certain.

Early demand faces a number of perceptual and human factor barriers, such as fears of invasion of privacy through centralized vehicle-highway interfacing, demonstrated reliability and safety, willingness to release control to a computer, uncertain benefits, and high initial cost (*7*). Credibility is especially an issue with the AVCS. Elias et al. (*8*) believe that an automated freeway system must be 20 times more safe than the current system to be acceptable to the public because of the driver's loss of direct control over the vehicle. Johnston et al. (*2*) note that commercial air travel is about 10 times as safe as automobile travel but that the public is more concerned about deaths from air travel than about deaths from automobile travel because of the large accidents and involuntary risk in air travel.

There are currently no public studies of consumer willingness to purchase driving automation. However, planners will need an advanced understanding of consumer responses. Because most AVCS technologies are still in development, consumer studies are in the exploratory stage, investigating potential early adopters and those with driving experiences related to proposed technologies (*9*).

METHOD AND RESEARCH DESIGN

This study was designed to investigate consumer responses to two AVCSs in advanced stages of development: (a) adaptive cruise control (ACC) and (b) the collision avoidance system (CAS). ACC is a radar-controlled system for adapting the throttle speed of a vehicle to that of a vehicle in its path. The CAS is a radar-controlled braking device that assumes control of braking when objects are detected in a car's pathway and the driver does not respond.

It was hypothesized that avid users of cruise control are potential early adopters of ACC and CAS technologies, with the assumption that these drivers are more willing to adopt automated driving technologies for their convenience. This hypothesis was formulated from innovation theories that suggest that early adopters of a technology are likely to be early adopters of subsequent related technologies.

The choice to use focus groups was based on the need to probe consumer responses to a radical shift in technology without the benefits of actual product testing. Four focus group interviews were conducted, each with 11 participants. Two groups were composed of avid users, and the other two of infrequent users. Infrequent users of cruise control were

Institute of Transportation Studies, University of California, Davis, Calif. 95616.

the control group. Ownership of cruise control could not be used as an identifying characteristic because consumers must often purchase a bundle of power options to get only one of the options. At least 3 participants confirmed this hunch, stating that they had purchased the power options bundle to get power mirrors and were not interested in cruise control.

Participants were recruited from an automotive market research data base that contained 8,000 households in Santa Clara, California—a congested urban area just south of San Francisco. Participants were recruited through phone interviews in which they were asked whether they owned a recent-vintage car with cruise control and how frequently they used that feature. Avid users were defined as respondents who liked using cruise control and used it every week. Infrequent users owned but generally disliked cruise control and used it less than weekly. Participants were paid $40 to participate.

The selection process resulted in a mixture of middle-class professionals, homemakers, and retired persons, with a high number of engineers (three to four per group)—the result of Santa Clara's computer manufacturing economy. Two-thirds of the avid group were males, and two-thirds of the infrequent users were females. Avid users reported higher annual miles traveled than infrequent users.

Cruise controls owned by the participants were of several types. The controls are located on a lever extending from the steering column, on the dashboard, or on the front surface of the steering wheel. In addition, speed controls varied from those that have only an on-off function, to those with a step function. Several participants expressed greater satisfaction with the steering-wheel type of cruise control, which also included accelerator and decelerator buttons. They liked the convenient location and often used cruise control instead of the foot pedal for acceleration.

The group interviews were divided into three parts: a discussion of current cruise control, the concept of ACC, and the concept of the CAS.

USES OF CRUISE CONTROL

The most cited advantage of cruise control was the ability to relax from the task of controlling speed on long journeys (cited by 23 of 44 participants). The second most cited advantage was using cruise control to control speeding on long stretches, where it is possible to lose track of speed, and in urban situations with well-known speed traps. The third most cited advantage was fuel savings. Several other advantages were cited by one or two participants, including one person whose college friends had used cruise control to control speed after drinking.

The most cited disadvantage of cruise control was that it was difficult to use in local urban driving conditions and, therefore, of limited practicality (cited by 11 of 44 participants). The second most cited disadvantage of cruise control was that the driver might relax too much when the cruise control is in operation, not be alert to danger, or not have the foot in position to brake. Although seemingly inconsistent with the already stated benefits of being able to relax, discussion centered around the idea that cruise control should be used carefully and when appropriate. Whether the driver is relaxing too much was considered to be a matter of individual judgment. Several participants reported having seen other drivers resting their feet on the seat, dashboard, and even out the window while driving sections of open freeway and thought that was too relaxed. Several infrequent users were afraid to use cruise control for fear it would malfunction, or because it required them to look down to set it.

The participants revealed that cruise control can be used as a convenience and as a tool; as a convenience it allows them to relax from work, whereas as a tool it increases work efficiency. Although infrequent users stated that they generally dislike using cruise control as a tool in traffic, avid users reported using cruise control often as a tool in what they described as congested situations. All avid users and several infrequent users said they used cruise control to keep up with the flow of traffic on multilane freeways where traffic flows at a steady speed.

Several avid users said they often use cruise control to travel in groups of vehicles moving in the same lane at or above the speed limit on long stretches of open freeways. Caravans (platoon-like groups) are unspoken agreements among drivers to travel together at a speed set by a lead driver. Participants believed cruise control owners group this way on long highways, coordinating their speeds. Some drivers say they reduce headway by concentrating on the brakelights of the lead car instead of the one directly in front, thereby anticipating speed changes as many as 10 cars in advance. Only one participant in the avid users group said he avoids these caravans because the close headways seem risky.

Several avid users said they attempted to use cruise control in local urban driving whenever possible. These drivers were willing to reset more often and to switch lanes to avoid resetting; a few admitted to waiting until the last moment to avoid resetting. Owners of variable speed cruise control use their accelerate and coast buttons to adapt to cars in front. One described using his cruise control in this context as though it were a video game, the goal being not to reset. A special use of cruise control among avid users was to control speed on expressways with timed intersections. The final major occasion for use of cruise control was to control speed for known speed traps.

The convenience benefits of cruise control generated much less enthusiasm because of a sense that relaxed driving is not compatible with safe driving. All participants thought that the convenience use of cruise control should be restricted to open stretches of highway; avid users were adament that drivers should not use cruise control to relax in other driving contexts. What determined when and where a driver used cruise control depended on the judgment and skills of the driver. Although cruise control has no decision-making components to its design, two future products addressed in the remaining paragraphs respond to traffic conditions and vehicle interactions.

RESPONSES TO ACC CONCEPT

After discussing cruise control, the groups were presented with the ACC concept. ACC was described as a radar-based speed control system, which the consumer could set for a desired speed, with the radar system adjusting vehicle speed to match that of vehicles directly in front. Once the front car moves into another lane, the consumer's car would resume its set speed. In addition, the system uses an audible signal to warn the driver of upcoming vehicles in its path.

The primary response among avid and infrequent users of cruise control was questioning whether the system would malfunction. In particular, several respondents were concerned that radar was not a reliable technology. One respondent mentioned a story about police using radar and clocking a tree at 70 mph. Several participants thought radar would be less responsive to nonmetallic objects. Another thought the radar would have to "see way out there." Several wondered how well the beam could focus on what was in front without getting confused about things on the side. In the final two interviews, the word "sensor" was substituted for radar, raising less negative responses. In addition, respondents were concerned about dependence on the technology.

Parents in the groups had a mixed reaction to the device, fearing on the one hand that their teenage children would not have the proper judgment to know when to use the device and on the other hand thinking that the warning system would be a good device to teach children not to tailgate.

The warning system was seen in a favorable light as a safety supplement—to keep drivers from tailgating and to catch them when not alert—and as a way to instruct new drivers of safe driving habits. In fact, many of the infrequent users thought the warning system was worthwhile but that the ACC would not be of much use. Among avid users there was some interest in the warning system, but the adaptive mechanism was not seen as a major improvement over existing cruise control. Those who owned cruise controls with accelerator buttons noted that their use of cruise control was much like ACC. They recognized the utility of ACC but did not see any great advantages over the system they already possessed.

Among avid users there was a humorous but telling suggestion that what was really needed was a device to control the speed of the driver in front; the problem is seen as that of getting around the car ahead to maintain speed. These avid users were not enthusiastic about putting speed control in the hands of the slowest driver in a lane. In all groups there was a mostly accurate perception that such a technology might improve traffic flow but, to be effective, would have to be on all cars, and would therefore be standard rather than optional.

Participants were asked to estimate, by secret ballot, the price of an ACC system and what they were willing to pay. Participants were not happy to make such an estimate, because many had no idea what their current cruise control had cost. Price estimates for the ACC ranged from $300 to $2,800, with an average of $900. Twenty-three said they would not purchase ACC technology at all, whereas 21 indicated they would be willing to pay between $100 and $1,400. Drivers consistently exhibited a willingness to pay estimates that were half of their price estimates. These figures can be interpreted in two ways. On the one hand, they reflect relative disinterest, because even those that were interested said they would be willing to pay only about half of what they think ACC will cost. On the other hand, many were willing to pay significant amounts.

RESPONSE TO CAS CONCEPT

The CAS was described as a sensor-based system that warns drivers of collisions and applies brakes automatically if a collision is imminent and the driver is not taking corrective action. Response to the CAS was more emotional and engaged than that for ACC. The immediate response of participants in all groups was a concern about technical reliability. Respondents asked many questions about how the device would respond in normal and extreme conditions. There was a general perception that the device would not distinguish between dangerous and nondangerous objects in the road and objects coming from the side.

A further response was a concern that collision avoidance was a complex decision, that complex technology would be subject to failure under many circumstances, and therefore that CAS technology could not be relied on. Avid users dwelled on comparing their own skills with the machine. Several comments suggest that avid users thought the CAS would interfere directly with their own good habits. It was thought that the device would be particularly suited for avoiding rear-end collisions in bumper-to-bumper traffic. Participants cited their own lapse of attention under those conditions. Two participants had been rear-ended in freeway traffic by someone who did not see them. They both said it would have been great if the other driver had a CAS. Avid users (who tended to be most confident of their own skills) and a few nonusers who feared losing control to the technology were most opinionated about machine failure.

As with cruise control and ACC, participants debated the compatibility of relaxation and good driving goals. Several participants were concerned that drivers would learn to rely on the system instead of on their own instincts, leading to less than vigilant driving. Other participants debated the previous concerns, stating that the CAS could be an aid to new drivers, a safety edge in poor driving conditions, and a backup when attention is averted. The groups were interested in the safety and tool benefits of the CAS but were uneasy with the intended and possibly unintended use of the CAS as a convenience.

Participants estimated that the price of CAS products would be $500 to $3,000, with the average being $1,500. Fourteen participants said they would not purchase it, whereas 30 stated they would be willing to pay $150 to $1,800.

SUMMARY

Given the conditions under which they do most of their driving, participants in the focus groups placed more value and importance on the tool uses of vehicle control technologies than on convenience uses. They especially valued safety benefits, although they doubted the reliability of these technologies (especially when based on radar) and their ability to respond to different types of dangers and obstacles.

As a result of this greater interest in safety, more drivers indicated an interest in purchasing CAS products than ACC, although the amount that drivers were willing to pay, of those interested, was about the same for both technologies.

These findings are tentative and need to be confirmed in larger studies, but they suggest that the designers and marketers of IVHS technology would be advised to focus more on safety benefits than on convenience benefits and to be highly sensitive to drivers' reluctance to defer control to machines.

Moreover, it appears that avid cruise control users, because they are more interested in driving efficiency than safety or even convenience and believe their driving abilities to be superior to machine decisions, are not an attractive early market for ACC and CAS products. Infrequent users are a better market because of their belief in their own fallibility and therefore greater interest in the emergency assistance offered by warning signals and the CAS. Because convenience was considered to be an inappropriate use under current highway conditions, it is suggested that the earliest market for vehicle control technologies will be sensor-based information, such as audible collision warnings, to assist drivers. The second market will likely be for backup or shadow CAS devices, possibly becoming standard safety equipment. Only then, and with the development of exclusive access lanes for smart vehicles, are large numbers of consumers likely to purchase automated control for vehicles.

ACKNOWLEDGMENTS

This work was funded by the Program on Advanced Technology for the Highway (PATH) of the University of California, in cooperation with the State of California Business Transportation and Housing Agency, Department of Transportation. Appreciation is also extended to the participants in the focus groups.

REFERENCES

1. L. Saxton. Automated Highway System Considerations for Success. *Proc., 30th Annual Conference of the IEEE Vehicular Technology Society*, Dearborn, Mich., 1980.
2. R. Johnson, M. Deluchi, D. Sperling, and P. Craig. Automating Urban Freeways: Policy Research Agenda. *Journal of Transportation Engineering*, Vol. 116, No. 4, 1990, pp. 442–460.
3. K. Chen and R. Ervin. Developing a Research Program in Intelligent Vehicle-Highway Systems. In *Vehicle/Highway Automation: Technology and Policy Issues* (Appleby et al., eds.). SP-791. SAE, Warrendale, Pa., 1989, pp. 41–50.
4. *Report to Congress on Intelligent Vehicle-Highway Systems*. Report DOT-P-37-90-1. Office of the Secretary of Transportation, March 1990.
5. S. Shladover. Research Needs in Roadway Automation Technology. In *Vehicle/Highway Automation*, SAE, Warrendale, Pa., 1989, pp. 89–104.
6. R. W. Marans. User Acceptance of Intelligent Vehicle-Highway Systems (IVHS): Directions for Future Research. In *Automated Highway/Intelligent Vehicle Systems: Technology and Socioeconomic Aspects*. SP-833. SAE, Warrendale, Pa., 1990, pp. 97–102.
7. J. Vostrez. Vehicle/Highway Automation: Policy Issues and Strategies in California. In *Vehicle/Highway Automation: Technology and Policy Issues* (Appleby et al., eds.). SP-791. SAE, Warrendale, Pa., 1989, pp. 69–80.
8. J. Elias, D. Stuart, L. Sweet, and A. Kornhauser. *Practicality of Automated Highway Systems*, Vol. 39–41. Report SHWA-RD-79. 1979.
9. J. Nystuen. Framework for Assessing Travel Behavior Response to Intelligent Vehicle Highway Systems. In *Automated Highway/Intelligent Vehicle Systems: Technology and Socioeconomic Aspects*. SP-833. SAE, Warrendale, Pa., 1990, pp. 103–112.

Publication of this paper sponsored by Committee on Vehicle User Characteristics.

Factors Influencing Commuters' En Route Diversion Behavior in Response to Delay

ASAD J. KHATTAK, JOSEPH L. SCHOFER, AND FRANK S. KOPPELMAN

An understanding of drivers' en route decisions may help design strategies for ameliorating traffic congestion. A survey of downtown Chicago automobile commuters was conducted to investigate en route diversion in response to incident-induced congestion. The effects of factors such as source of congestion information (radio traffic reports versus observation), driver and trip characteristics, route attributes, and environmental conditions on driver response to delay were explored. En route diversion behavior was found to be influenced by source of traffic information, expected length of delay, regular travel time on the usual route, number of alternate routes used recently, anticipated congestion level on the alternate route, gender of the driver, residential location, self-evaluation statements about risk behavior (personality), and stated preferences about diverting. The results show that real-time traffic information broadcasts provide a basis for en route diversion decisions. Further, length of delay and perception of traffic congestion on the alternate route also influence such decisions. Short-term improvements in real-time traffic information should focus on disseminating information about length of delay due to incidents and the congestion levels on the alternate routes surrounding the incident. This action requires monitoring traffic conditions on the alternate routes along major roadways. Providing clearer information on delays and congestion will help drivers make more informed route selection decisions.

Traffic congestion occurs when the vehicles using a roadway impede each other's progress. The consequences of congestion include delays, accidents, excessive fuel consumption, air pollution, and driver frustration. One way to ameliorate traffic congestion is through demand management strategies designed to modify driver behavior by encouraging drivers to change modes, routes, and departure times.

Congestion can be incident induced or recurring. Incident-induced congestion may be caused by accidents, vehicle disabilities, short-term maintenance activities, weather, and so on. Recurring congestion is caused by an increase in demand during rush hours. McDermott (*1*) has estimated that each causes about half of the congestion on Chicago area freeways. The congestion problem becomes acute when incident-induced congestion occurs during recurring congestion (peak period).

One demand management strategy used in larger cities is the collection, processing, and dissemination of near real-time information about traffic conditions on key highway links. This information usually includes qualitative or quantitative descriptions of congestion, reports of incidents, and, in the Chicago area, estimates of point-to-point travel times. Broadcast traffic information in Chicago is relatively sophisticated.

Civil Engineering Department and The Transportation Center, Northwestern University, 1936 Sheridan Road, Evanston, Ill. 60208.

Real-time traffic information on the downtown-oriented freeways is gathered with a network of loop detectors maintained by the Illinois Department of Transportation (IDOT). Many radio stations compete for the drive-time market by providing rush-hour traffic reports: some use IDOT information alone, and others buy traffic information from specialty firms that mix reports from ground and aerial observers with IDOT information. A few radio stations supplement these sources with their own networks of observers.

Real-time traffic information is intended to help travelers make better choices of modes, departure times, and routes, as well as to reduce their en route anxiety. Considerable public and private resources are devoted to collecting and disseminating this information, so it is important to understand its impact on motorists. As congestion grows and there are fewer capital-intensive options to ameliorate it, the potential importance of near real-time traffic information grows. The worldwide interest in the intelligent vehicle-highway system (IVHS) is founded on assumptions about driver responses to such information (*2*).

An understanding of drivers' en route decisions is important for designing congestion-reduction strategies, particularly the IVHS. Relatively little is known about such decisions, and this study was conducted to contribute to that body of knowledge, specifically to investigate factors influencing en route response to incident-induced congestion. The effects of real-time traffic information along with factors such as driver, trip, and roadway characteristics were evaluated.

Several researchers have found that a significant number of drivers divert in response to delay on their preferred route (*3–7*). Therefore, delay was selected as the criterion for investigating driver response to incident-induced congestion. It was decided to interview downtown Chicago automobile commuters because the Chicago area traffic monitoring system is downtown focused and because downtown automobile commuters regularly experience congestion.

CONCEPTUAL FRAMEWORK

It was hypothesized that drivers are influenced by real-time traffic information, such as traffic reports, and by the following factors (see Figure 1):

- Incident characteristics, such as length of delay;
- Trip characteristics, such as trip origin and destination and availability of alternate routes;
- Attributes of preferred and alternate routes, such as travel time and scenery;

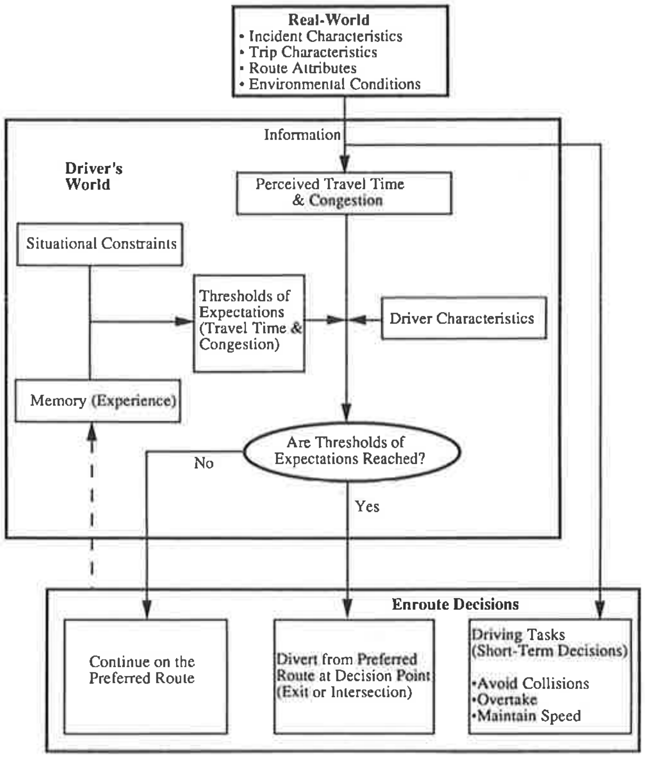


FIGURE 1 Conceptualization of en route decision-making process.

- Environmental conditions, such as weather;
- Driver characteristics, such as age, gender, and personality;
- Work rules, such as flexibility in work arrival time and type of work; and
- Situational constraints, such as remaining trip length.

Drivers perceive information through direct contact with the environment; in response they perform driving tasks; that is, short-term decisions such as overtaking and maintaining speed. Commuters making regular work trips compare the perceived travel time and congestion to their expectations. If thresholds of their expectations are reached (when there are large differences between perceptions and expectations, e.g., when there is a major delay), the driver experiences frustration and may be prompted to make a diversion decision. The thresholds of expectations may vary for the same person depending on the situation and, of course, may vary among individuals. In some cases, a driver may be frustrated with a delay situation but continue on the selected route. Although it may be impossible to measure time and congestion thresholds, revealed behavior can indicate if the thresholds were reached: if a person diverts in response to delay, the individual's time threshold was reached.

RELEVANT LITERATURE

Diversion behavior of drivers has been investigated primarily in the context of short-term maintenance operations and special events. The method used can be categorized as (a) the stated preference approach, (b) the revealed (or reported) preference approach, and (c) the field study approach. The stated preference approach analyzes driver predictions of their behavior in response to hypothetical scenarios; for example, respondents might be asked if they would divert in response to a particular delay. The revealed preference approach analyzes drivers behavior in real-life situations on the basis of respondents' reports about previous actions; for example, respondents might be asked if they diverted in a specific case. The field study approach analyzes driver behavior through field observations of drivers, for example, observation of actual diversion behavior in response to specific incidents.

Self-predicted behavior of respondents may be quite different from their revealed behavior, and thus it is not uncommon to find differences between stated and revealed behavior. The revealed preference approach was used, although stated preferences were used to a lesser extent for added insights and for comparison with revealed preferences.

Table 1 presents a summary of factors that influence diversion behavior, along with the research approach used in the study. Overall, the factors found to induce diversion were traffic information (*3,4,7–14,18,20,21*), longer travel time on the preferred route (*5–7,11*), congestion and delay on the preferred route (*3–6,8–11,13,18,19*), and familiarity with the alternate route (*10,11*). Factors found to inhibit diversion were longer travel times on the alternate route (*5,6,11,18*), expected congestion and delay on the alternate route (*8,11,18*), and traffic stops on the alternate route (*5,18*). Young, male, and unmarried drivers had a higher inclination to switch routes (*6,7*).

Using the stated preference approach, Huchingson and Dudek (*4*) and Huchingson et al. (*9*) found the relationship between delay and diversion to resemble an S-shaped curve, with few drivers expressing willingness to divert in response to minor delays and all but about 5 percent expressing willingness to divert in response to 30- to 60-min delays. In contrast, researchers using the revealed preference approach have found that a large portion of the driver population did not divert at all. For example, Huchingson et al. (*11*) found that only 60 percent of commuters in Dallas had taken one or more alternate routes; Daniels et al. (*12*) found that 36 percent of Chicago drivers surveyed had never diverted. Thus, drivers may express willingness to divert in hypothetical situations, but their actual diversions may be considerably less, perhaps influenced by a variety of situational variables.

Huchingson and Dudek (*4*) used the stated preference approach and found that the median value of delay for diversion was 15 to 20 min for various groups from different locations in the United States. A similar study in Houston found the median value of delay for diversion to be 5 to 6 min, which might be due to the availability of a convenient service road as an alternate route (*9*). Haselkorn et al. (*10*) found that the averages of length of delay before diverting vary between 13.5 and 27.4 min.

Several researchers using the revealed preference approach have found that a significant number of drivers divert from their preferred route. For example, Heathington et al. (*8*) found that on the average commuters diverted 23 percent of the time due to congestion on their preferred route. Huchingson et al. (*11*) found that 27 percent of the respondents diverted in response to incidents (which they were asked to recall), whereas 69 percent continued on their preferred route.

TABLE 1 SUMMARY OF FACTORS FOUND TO INFLUENCE DIVERSION BEHAVIOR (SELECTED STUDIES)

Author(s)	Traffic Information	Travel Time Preferred Route	Travel Time Alternate Route	Delay on Preferred Route	Congestion on Preferred Route	Congestion on Alternate Route	Familiarity with Alternate Route	Traffic Stops on the Routes	Location Sample Size & Methodology [1]
Heathington, et al. (*8*)	+			+		-			Chicago, IL. N=732. SP, RP
Dudek, et al. (*3*)	+				+				Houston & Dallas, TX. N=505. RP
Haefner & Dickinson (*30*)		+							Baltimore-Washington. N=? RP
Daniels, et al. (*12*)	+								Chicago, IL. N=732. RP
Huchingson, et al. (*11*)	+	+	-	+	+	-	+		Dallas, TX. N=202, 215. RP
Richards, et al. (*18*)	+		-	+		-	+ -	-	Dallas, TX. N=480. RP, FS
Turner, et al. (*21*)	+								Chicago, IL. FS
Dudek, et al. (*14*)	+								Dallas, TX. FS
Huchingson & Dudek (*4*)	+			+					Multistate. N= 240, 184, 40. SP
Dudek, et al. (*19*)	?			+	+				San Antonio, TX. FS
Roper, et al. (*20*)	+								Los Angeles, CA. FS
Huchingson, et al. (*9*)	+			+					Houston, TX. N=843. SP
Shirazi, et al. (*13*)	+				+				Los Angeles, CA. N=400. RP
Stephanedes, et al. (*5*)		+	-	+				-	Twin Cities, MN. N=105, 195. RP
Mannering (*6*)		+	-		+				Seattle, WA. N=117. RP
Haselkorn, et al. (*10*)	+			+	+		+		Seattle, WA. N=3893. RP, SP
Mahmassani, et al. (*7*)	+	+ -							Austin, TX. N=372. RP

+ = Induces Diversion; - = Inhibits Diversion

[1] SP=Stated Preference approach (hypothetical scenarios).
RP=Revealed Preference approach (real-life situations described by respondents)
FS=Field Study approach (observation of driver behavior by researchers)

Shirazi et al. (*13*) found that 40 percent of the commuters in Los Angeles diverted on their way to work, and 14 percent of the respondents diverted either very often or often.

Traffic information has been found to influence diversion. Heathington et al. (*8*) found that driving patterns of frequent diverters were influenced slightly more by traffic reports than by visual observation. Mahmassani et al. (*7*, p. 11) found that drivers who generally listen to radio traffic reports have a greater propensity to switch routes. Daniels et al. (*12*) found that the average frequency of diversion for Chicago drivers was 16 to 27 percent for observed congestion and 17 to 35 percent for traffic reports. Thirty percent of the drivers in the Sharazi et al. (*13*) study said that radio traffic reports help them in their decision to divert.

For content and format of real-time traffic information, guidelines for highway advisory radio messages have been developed (*14,15*). Moreover, Dudek et al. (*16*) found from a laboratory study that drivers preferred terse messages rather than conversational style. Unfamiliar drivers going to a special event were more likely to divert in response to highway advisory radio diversion messages than were familiar drivers (*17*). Generally, studies that have used the revealed preference approach indicate that information either in the form of diversion advice or travel times and congestion can induce drivers to divert.

Field studies of diversion behavior have shown mixed results. Dudek et al. (*19*) evaluated the effect of changeable message signs on diversion in San Antonio, Texas, and could not find statistical evidence for increased diversion due to messages displayed under incident conditions. Roper et al. (*20*) found that drivers in Los Angeles were successfully diverted from a busy freeway during repairs through a comprehensive public management campaign and the use of changeable message signs during the closure. Diversion messages developed in laboratory studies were effective in diverting freeway traffic going to special events (*14,18*). Turner et al. (*21*) also found evidence of increased diversions due to changeable message signs; however, the findings were not based on statistically sound data bases for comparisons. Overall, field studies indicate that drivers can be successfully diverted to alternate routes during special events through diversion messages, but there is not enough evidence to suggest that drivers can be successfully diverted during incidents.

Several researchers have found that drivers were more likely to divert to familiar routes (*10,11,18*), suggesting that cognitive maps of drivers may influence their diversion behavior. Cognitive maps are stored information about the relative location of objects in the physical environment, and they have been linked to lifestyle, age, stage in family life cycle, and social class (*22*). The state of an individual's cognitive map indicates the level of familiarity with the network and route alternatives. The number of alternate routes used by a driver was chosen as a proxy for cognitive maps. Freundschuh (*23*) provides a review of cognitive maps in the context of developing IVHS technology.

METHODOLOGY

Sampling

The target population for empirical work consisted of automobile drivers who made repeated trips during which broadcast traffic information was available to them. Drivers who made regular trips were of interest because it is easier for the respondents and the researchers to identify behavior patterns, and it is probably these drivers who experience the worst congestion on a regular basis. The importance of the availability of traffic information is obvious, because it is the effect of this factor that was to be explored.

Due to the radial orientation of the freeway network in Chicago and the availability of several alternate routes, it was decided to sample work trip drivers destined to the central business district (CBD). These travelers were intercepted at a sample of downtown parking garages during rush hours (7:00 to 10:00 a.m.). Mail-back questionnaires were distributed to commuters as they walked out of garages in the morning during April 1990. A total of 700 questionnaires was received, representing a response rate of 33 percent.

Self-selection bias is inherent in surveys of this type. Some degree of bias may be a reasonable price to pay for obtaining insights into real-world diversion behavior and for refining hypotheses. To check for obvious biases, the socioeconomic attributes and trip characteristics of this sample were compared with those of other samples of commuters who drove to downtown Chicago (*24*). This sample compared reasonably well with similar samples of commuters in the Chicago area, and it was consistent with expectations about downtown Chicago commuters (*25*). There were no indications of substantial distortions in the socioeconomic characteristics of the respondents, which does not preclude the possibility of self-selection bias in other dimensions (e.g., behavioral patterns, personalities, and cognitive maps).

Instrument Design

Driver response to a specific, recent delay experience was investigated as opposed to delays in general because this factor was expected to produce more meaningful results. First, the respondents were asked if they knew about an en route delay longer than 10 min on their way to or from work during the past 6 months. There was a possibility of memory loss or distortion of responses to the delay experience because of the length of this time interval. However, the data showed that, among the 62.5 percent of respondents who knew about a delay longer than 10 min, 77.5 percent had experienced this delay within the past 2 months. Therefore, the influence of memory loss and distortion seems small for a majority of the respondents. The survey then asked about the details of the delay experience, for example, how and when drivers received the delay information. At the end of the section on delays, the survey asked if respondents diverted to an alternate route because of the delay.

The usual and alternate routes were defined as substantially different to avoid the complexity of dealing with a large number of small variations in the route of travel. Furthermore, there is no evidence in the literature to indicate that drivers consider a large number of overlapping alternatives when selecting routes. Stephanedes et al. (*5*) found that less than 3 percent of the commuters in Minneapolis–St. Paul considered more than two alternatives, and, if they did, it was under unusual circumstances such as a blizzard. It was explained to the respondents through an illustration that the predominant roadway on the alternate route should be different from that on the usual route. To further clarify the definition, respondents were given an example in which a freeway was the usual route and a parallel arterial street was the alternate.

The Chicago area experienced a severe snowstorm on February 14, 1990 (known as the Valentine's Day snowstorm), with disastrous consequences for automobile commuters. Results of this study showed that the average travel time on the trip home was 163 min instead of 46 min under normal conditions. There was concern that experiences on the day of this blizzard might distort responses to the survey. Because the objective was to investigate delays caused by day-to-day congestion and not delays due to weather extremes, the questionnaire included a separate section on the blizzard (results not reported here) to separate routine experiences from unique circumstances.

OVERVIEW OF RESPONSES

A profile of the respondents' socioeconomic attributes is presented in Table 2. The sample represents a stable, upper income, well-educated, and well-established group. One of the objectives was to sample regular peak-period automobile commuters, and Table 3 indicates that this goal was accomplished. Most respondents traveled regularly to downtown Chicago; about 86 percent had traveled more than 10 times to downtown Chicago during the past month.

Most respondents (93 percent) started their work between 7:30 and 9:00 a.m. The median work start time was 8:30 a.m. More than 90 percent of the respondents could not arrive more than 1 hr late "without it mattering much." The mode and median for arrival time flexibility were both 15 min. Most

TABLE 2 SUMMARY OF RESPONDENTS' SOCIOECONOMIC ATTRIBUTES

Sample Attributes		Frequency %	Sample Attributes		Frequency %
Age	20-29 Years	23.6	**Household Size**	1 Person	16.6
	39-39 Years	33.0		2 persons	35.5
	40-49 Years	29.7		3 persons	19.5
	50-65 Years	12.5		4 persons or more	28.4
	65 Plus Years	1.2			
Gender	Male	54.3	**Area of Residence**	Suburbs	52.7
	Female	45.7		Chicago	47.3
Occupation	Managerial/Business owner	23.7	**Time at Present Job Location**	Up to 1 Year	13.9
	Clerical	7.7		Between 1-2 Years	19.1
	Professional/Technical	54.5		Between 3-5 Years	23.3
	Sales	9.2		Between 6-10 Years	17.5
	Other	4.9		More than 10 Years	26.2
Education	High School	4.6	**Time at Present Home Location**	Up to 1 Year	11.5
	Vocation/Technical School	2.0		Between 1-2 Years	21.2
	Some College	19.4		Between 3-5 Years	24.0
	Graduated College	34.1		Between 6-10 Years	13.9
	Post Graduate Work	39.9		More than 10 Years	29.4
Income	Up to $20,000	4.3			
	$20,000-$40,000	33.1			
	$40,000-$60,000	26.3			
	$60,000-$80,000	14.0			
	$80,000-$100,000	6.6			
	$100,000 Plus	15.7			

Percentages do not consider missing data. Total Sample Size=700.

TABLE 3 SUMMARY OF RESPONDENTS' TRIP CHARACTERISTICS

Sample Attributes		Frequency %
Number of downtown Trips During Past Month	Up to 10 Times	14.1
	11-15 Times	10.0
	16-20 Times	20.2
	21 Times or More	55.7
Work Start Time	Up to 8:00 AM	39.9
	Between 8:01-8:30 AM	36.8
	After 8:30 AM	23.3
Flexibility in Work Start Time	0-15 Minutes	53.2
	15-60 Minutes	38.1
	More than 60 Minutes	8.7
Car Occupancy	1 Person	67.6
	2 Persons	26.8
	3 or More Persons	5.6
When Route is Chosen	Before Getting in Car	74.0
	After Getting in Car	26.0
Years Used Usual Route	Up to 1 Year	20.2
	Between 1-4 Years	49.4
	More than 4 Years	30.4
Number of Alternate Routes Used	None	7.8
	1 Route	28.0
	2 Routes	32.7
	3 Routes	18.1
	4 or More Routes	13.4

Percentages do not consider missing data. Total Sample Size=700.

respondents chose their route before getting in their cars. Close to 80 percent had used their route for more than a year, although not without variation. Most respondents (63.2 percent) had taken two or more substantially different alternate routes between home and work. The average travel time to work was 42.6 min and that from work was 46.1 min. As expected, the best alternate route took longer on the average (49.8 min).

About 62 percent of the respondents experienced en route delays during the past 6 months (see Table 4). Nearly 85 percent expected the delay to add between 10 and 30 min to their work trip. A majority of the respondents (60.8 percent) described their delay experience on the home-to-work trip. Forty-three percent said they received the information about delays through radio traffic reports. More than 42 percent took an alternate route in response to en route delays. Because the alternate route was defined as substantially different from the preferred route, the diversion decision represents a major behavioral change.

To assess the influence of driver personality on en route diversion, 14 self-assessment questions were used. Some of these were borrowed from the psychology literature, as well as from other work that explored the travel behavior relationship (*26*). Several questions developed in this study were added, focused specifically on behaviors likely to describe route choice.

Several personality factors were identified using factor analysis. Only the most relevant factor, termed "adventure and discovery," is reported here. On the basis of the responses to the statements presented in Table 5, this factor represents respondents' propensity toward risk and exploration. It was expected that drivers who are risk prone and inclined toward exploration may be more willing to divert from their preferred route. The sums of respondents' scores on these statements were used to model diversion behavior.

Stated preferences of drivers about diversion were investigated. This appraoch was intended to enrich the perspective on diversion behavior and compare the responses with findings from the revealed preference approach. The respondents gave their preferences about diversion in hypothetical situations, and the frequencies are presented in Table 6. The sums of scores on these statements, which represent a driver's propensity to divert, were used to model diversion behavior.

TABLE 4 SUMMARY OF RESPONDENTS' DELAY EXPERIENCE

Questions	Frequency %	Number of respondents
Do you know of a traffic "delay" longer than 10 minutes after you got on your "usual" route during the past six (6) months?		
Yes (Experienced Delay)	62.5	437
No	37.5	262
How Long Ago did this happen?		
Less Than 1 Month	59.7	258
1-2 Months	17.8	77
2-3 Months	12.0	52
3-4 Months	5.6	24
More Than 4 Months	4.9	21
How much time did you expect the "delay" to add to your trip?		
10-20 Minutes	57.6	250
21-30 Minutes	27.0	117
31-40 Minutes	9.9	43
41-50 Minutes	3.9	17
More Than 50 Minutes	1.6	7
When you got the information about the "delay" were you on your way from:		
Home to Work	60.8	261
Work to Home	39.2	168
What was the weather like when you got the information on the "delay?"		
Clear	41.8	180
Cloudy and Dry	20.8	90
Rainy/Light Snow	35.3	152
Blizzard/Storm	2.1	9
How did you get the information on the "delay?"		
Radio Traffic Reports on Congestion Ahead	43.0	187
Your Own Observation of Unusually Heavy Congestion	55.4	241
Don't Recall	0.7	3
Other	0.9	4
Did you take an alternate route after getting the information on the "delay" on your usual route?		
Took Alternate Route	42.5	185
Continued on Usual Route	57.5	250

TABLE 5 ADVENTURE AND DISCOVERY FACTOR OF DRIVER PERSONALITY

Statement of the Question	Source of the Statement	Average Score
I like discovering new routes to get someplace	Modified from Ergun (*26*)	2.05
I sometimes do things just to see if I can	Ergun (*26*)	2.14
I am willing to take risks to avoid traffic delays	Modified from Ergun (*26*)	2.28
I like exploring new places	Modified from Ergun (*26*)	2.81
I am not afraid of getting lost in the Chicago area		2.89
I would rather take a little longer to use a route I know well		2.27

Responses coded as: 0=strongly disagree, 1=disagree, 2=neutral, 3=agree, 4=strongly agree.

IS THERE A USUAL ROUTE?

To understand diversion behavior, it is important to know whether drivers indeed have a regular, or base, route from which they occasionally divert. To answer this question, we analyzed how often respondents used their route, given their trip frequency, as well as the length of time respondents had used their route. Slightly more than 80 percent of the respondents had used the same route for more than 1 year, although about 70 percent had diverted from it occasionally. Drivers were significantly more likely (at the 5 percent level) to use their route for less than a year if they had worked at their current job location or lived at their current home address for less than a year. Furthermore, most drivers took their route regularly, as shown by a high correlation between number of trips and number of times the route was taken (correlation coefficient = 0.88). Respondents over 40 years of age were more likely to have used their route for more than 4 years and, in 3.5 percent of the cases, for more than 20 years. These results imply that most drivers stick to one route as opposed to constantly shopping for new routes. Therefore, it is inferred that there is a usual route that drivers use most frequently and for a longer period of time. At the same time, drivers in this sample did switch routes occasionally.

DELAY EXPERIENCE

More than 62 percent of the drivers knew about an en route delay on their usual route. This figure seemed lower than expected. The concern was that length of the questionnaire (8 pages, 112 questions) might have affected the response to this section; moreover, respondents who skipped this section to save time might introduce a bias because they may have a higher value of time. To address this concern, the following analysis was conducted.

Cross tabulation between the response to whether the respondent knew about a delay and income or education did not show a significant relationship (5 percent level). Further, it was expected (and confirmed) that drivers who experienced longer travel times would be more likely to know about en route delays. Suburban residents were more likely to know about a delay than were Chicago residents, as expected (5

TABLE 6 PERCENTAGE DISTRIBUTION OF RESPONSES TO STATED PREFERENCE QUESTIONS

Conditions on Usual Route	Conditions on Best Alternate Route	Definitely Take "Usual" Route		Neutral		Definitely Take "Alternate" Route	Sample Size
15 minute delay on trip **to work and you must be on time** to work today	No delay from "normal" conditions	25.6	12.7	11.6	14.2	35.9	640
15 minute delay	**No delay** and **no traffic signals/stop signs** on route	22.4	12.2	14.0	15.9	35.5	630
15 minute delay predicted by your own **observation of heavy congestion**	No information on current conditions	29.7	23.0	19.2	16.6	11.5	640
Jammed (Stop and Go)	No delay from "normal" conditions	12.2	9.1	14.2	21.6	42.9	639
15 minute delay due to **heavy traffic**	No delay from "normal" conditions	19.1	20.0	16.0	18.5	26.4	639
15 minute delay due to **accident**	No delay from "normal" conditions	15.3	15.0	13.3	22.0	34.4	640
15 minute delay **reported on radio**	No information on current conditions	22.6	21.2	25.1	14.8	16.3	637
15 minute delay	No delay from "normal" conditions	21.3	18.2	21.6	16.7	22.2	633
15 minute delay on return trip **from work**	No delay from "normal" conditions	26.7	18.6	15.3	16.7	22.7	633
15 minute delay	No delay but **one traffic signal every one-half mile** of the route	33.9	24.8	19.3	12.8	9.2	632
20 minute delay	No delay from "normal" conditions	15.3	9.5	13.7	22.5	39.0	634
10 minute delay	No delay from "normal" conditions	42.2	21.4	16.7	8.0	11.7	635

percent level). Respondents who made more regular trips and drivers who started work between 8:00 and 8:30 a.m. were more likely to know about a delay. Males were also more likely to know about a delay. Overall, the distribution of negative responses to the delay question seemed reasonable, and there was no obvious discrepancy. Any underreporting to save effort on the survey appears to be uniform across groups.

Length of Delay and Diversion

Several researchers have found that length of delay on the preferred route influences behavior (*3–10*). Increasing delays on the preferred route cause more drivers to divert, and the relationship between length of delay and diversion derived with the stated preference approach resembles an S-shaped curve.

The relationship between length of expected delay and diversion to an alternate route was investigated (see Figure 2). Relatively fewer drivers diverted to alternate routes in response to an expected delay of 10 to 20 min and, as expected, the percentages of drivers diverting increased when the delay increased to 21 to 30 min. However, among drivers who expected the delay to add more than 30 min, the percentage of drivers diverting did not increase. In fact, among the drivers who expected the delay to add 40 min or more to their work trip, only 42 percent diverted. This result is counterintuitive, and possible reasons for it are discussed in the following paragraphs.

The stated preference approach asks people to predict their reactions to delays of specific lengths. Because respondents know the delay length with certainty, it is easy for them to present a rational response (*4,9*). As shown in Figure 3, frequencies of responses to various delay intervals in the stated preference questions (see Table 6) indicate such a rational response. In real-life situations, however, drivers cannot usually know with certainty the length of delay in advance. In retrospect, a driver may feel that large amounts of time could have been saved by diverting; however, initially the delay may not have been threatening enough to divert. As a result, in real life, driver response may not display the same level of rationality. Thus, one reason for a decrease in the percent diverting for large delays may be that delay is perceived incrementally. Furthermore, there were only 24 observations in the more-than-40-min delay category, and this small sample may have contributed to the counterintuitive result.

Other variables, such as trip direction (to or from work) and weather conditions, might have explained why the percent diverting decreased at the upper end of delay distribution. For example, if long delays occur when the weather is particularly bad, no alternate route is likely to be more appealing. The relationship between delay and diversion was investigated while controlling for weather conditions; the results showed that bad weather (rain or snow) did not inhibit drivers from diverting to their alternate routes. In fact, the results suggest that, in bad weather, increasing delays cause the percentage of people diverting to increase.

The relationship between delay and diversion at the upper end of the distribution was unexpected, yet analysis of real-life situations can sometimes give unexpected results. This finding may reveal something about how drivers get information about delays, and it underscores the absence of predictive information about traffic conditions. The contrast between results from this study and those based on stated preferences suggests that more reliable information may lead to increased diversion. For example, the S-shaped response to stated preference questions might be interpreted as an indication of how drivers want to behave, and the irregular response to revealed preference questions might be interpreted as a description of how they must behave given the information they have available.

Saving and Loss in Travel Time

The distribution of saving and loss in travel times due to either diverting or staying on the usual route was investigated. Drivers who diverted were asked to give an estimate of travel time saved or lost by diverting. Drivers who did not divert were asked to give an estimate of the travel time that could have been saved or lost by taking the best alternate route. Of course, these reported time savings or losses are subjective driver estimates, which may be exaggerated consciously or

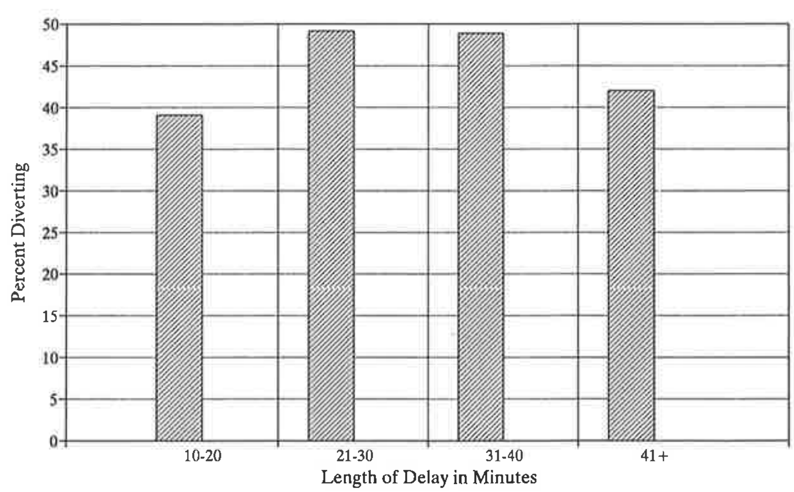


FIGURE 2 Relationship between diversion and length of delay.

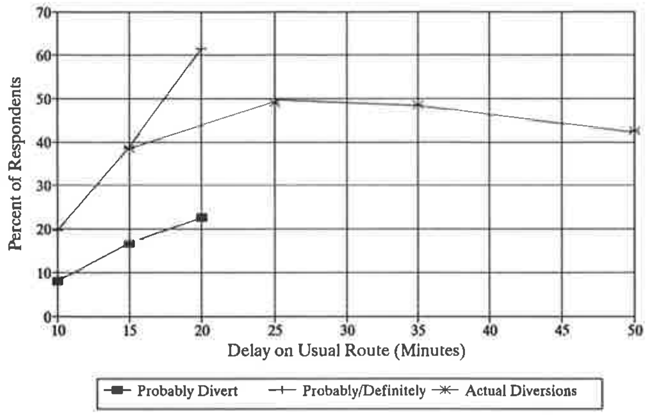


FIGURE 3 Stated intentions and revealed behavior as a function of delay.

subconsciously to rationalize their actions. The magnitudes of travel time saving or loss shown in Figure 4 may represent the travel time thresholds referred to in the conceptualization. The distribution in Figure 4 suggests that 23.1 percent of those who diverted believed that they lost travel time, whereas half of those who continued on their usual route believed that they would have saved time by taking their best alternate. This finding indicates that a substantial number of drivers stay on their usual route, with which they are more familiar, even in the face of a time loss. Furthermore, half of the respondents who believed that they could have saved time by diverting may not have done so because delay may be perceived incrementally. More than 76 percent of those who diverted thought that they gained time, and 26.6 percent estimated the saving to be more than 15 min. The two distributions shown in Figure 4 are bell-shaped. The average saving in travel time for those who diverted was 9.6 min, and the expected loss in travel time for those who did not divert was 4.2 min. Other researchers have found the median delay before diversion to vary between 5 and 27 min (*4,9,10*); therefore, the reported savings in travel times seem reasonable. Again, these savings and losses should be viewed with caution because drivers may exaggerate them to justify their decisions. Nonetheless, statistical analysis suggests that the results are reasonable and consistent with previous studies.

MODELING DRIVER RESPONSE TO DELAY

The effect of several variables on the diversion decision was examined by estimating diversion choice models on the basis of respondents' reported experience of a recent delay. To model choice using utility maximization, knowledge of the decision to be made, the alternatives and their attributes, and the individuals' attributes is needed (*27*). The decision is to divert (to an alternate route) or to stay on the usual route. Although there may be several alternate routes available, these choices were reduced to staying on the usual route and diverting to an alternate. Consideration of more than two alternatives would have added complexity for the respondents and the analysts. The attributes of usual and alternate routes were selected on the basis of research of commuter route choice (*11,28,29*). They included ratings on a five-point Likert scale ranging from "strongly agree" to "strongly disagree" for the following attributes: congestion, scenery, reliability, neighborhood safety, stress experienced while driving, traffic stops, and overall evaluation of the route. The individual attributes include socioeconomic and demographic characteristics, personality (from the self-assessment statements), and attitude toward diversion (from the stated preference questions).

The effects of the following variables were explored:

- Characteristics of the delay experience, such as weather, trip direction, length of delay, and information source on delay;
- Attributes of usual and alternate route;
- Trip characteristics, such as respondents' estimates of travel time on the usual and alternate routes and length of time the usual route had been used; and
- Socioeconomic attributes of the respondents, such as age, gender, income, and location of residence, as well as driver personality factors developed in this study.

Theoretical justification and statistical tests were used to choose among alternative models. Table 7 shows the selected logit model of diversion choice on the basis of these criteria. The signs and magnitudes of the coefficients are as expected. The base for the model was not diverting; therefore, a positive

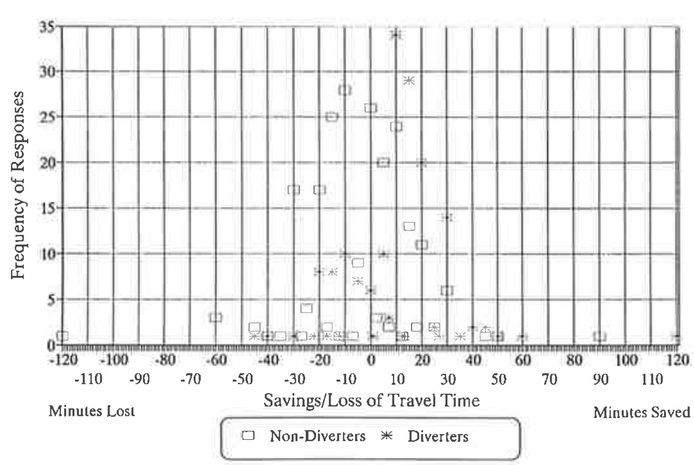


FIGURE 4 Perceived time savings and loss as a function of diversion behavior.

TABLE 7 MODEL OF DIVERSION CHOICE

Variables	Coefficients	(t-Statistics)
Constant	-6.056	(-3.02)
Information source (0=Radio Traffic Report on delay, 1=Observation of delay)	-0.649	(-2.28)
Number of *alternate* routes used (0=0 Routes, 1=1 Route, 2=2 Routes, 3=3 Routes, 4=4+ Routes)	0.303	(2.35)
Alternate route is congested (0=Strongly Disagree, 1=Disagree, 2=Neutral, 3=Agree, 4=Strongly Agree)	-0.243	(-2.00)
Logarithm of length of delay (Minutes)	0.619	(1.53)
Logarithm of Travel Time (Minutes)	0.676	(1.58)
Gender (0=Male, 1=Female)	-0.639	(-2.17)
Stated Preference Index (sum of the scores on stated preference questions normalized by number of statements)	0.682	(4.68)
Personality (sum of the scores on "Adventure and Discovery" normalized by number of statements)	0.412	(2.00)
Residence (1=North suburbs of Chicago, 0=Otherwise)	-0.409	(-1.08)
Residence (1=South suburbs of Chicago, 0=Otherwise)	-1.046	(-2.47)

Summary Statistics

Initial log-likelihood	-198.24
Log-likelihood at convergence	-160.19
Number of observations	286
Percent correctly predicted	71.33
Rho-squared	0.1919
Rho-squared bar	0.1364

sign indicates increased likelihood of diverting. The sign of the constant is negative, reflecting a preference for not diverting, which is expected because diversion is an unusual occurrence and, even if there is a delay on the usual route, drivers are expected to prefer to stay on it. The null hypothesis that each parameter value is zero can be rejected at the 5 percent significance level except for the delay, travel time, and one of the residential location variables.

Travel time, length of delay, and location of residence have low *t*-statistics; however, these variables are included in the model because longer travel times offer more opportunities for diversion, longer delays on the preferred route prompt drivers to divert, and suburbanites may have a different awareness of urban route alternatives or different levels of comfort with diversion in urban areas. Likelihood ratio tests indicated that none of these variables should be dropped.

It was expected that drivers would be more likely to divert if they observed traffic delays as opposed to receiving delay information through radio traffic reports. Yet the sign of information source is negative, which means that drivers were more likely to take alternate routes if they received delay information through traffic reports. This finding is consistent with those of Heathington et al. (*8*), who found that frequent diverters (due to congestion) were influenced slightly more by traffic reports than by visual observations. Furthermore, Mahmassani et al. (*7*) found that drivers who listened to radio traffic reports were more prone to divert.

Drivers may be more inclined to divert in response to traffic reports because they may have more options at the time they get radio information on incident-induced congestion. By the time drivers observe traffic congestion, they may have committed themselves to a route, or they may not have any real diversion options. The observed delay may be perceived in increments, which may further inhibit a driver from diverting, whereas traffic reports tend to give a more global picture of congestion. Overall, incident-related delay information has potential for modifying driver behavior.

Logarithmic transformation of length of delay and travel time variables was found to be statistically superior to the linear specification. This finding suggests a reduced sensitivity to units of delay and trip time increases; a given percentage increase in the length of delay has the same effect on diversion regardless of the current value of delay. For example, a 5-min increase in a 10-min delay (a 50 percent increase in the length of delay) has the same effect on driver decisions as adding 15 min to a 30-min delay. This finding is reasonable because a driver who anticipates experiencing a half-hour delay may not care that much about an additional 5-min delay, whereas a driver who anticipates a 10-min delay cares more about an additional 5-min delay. The logarithmic transformation seems applicable within reasonable limits of delay (of up to 1 hr). The positive signs indicate that, when delay on the usual route was higher, drivers were more likely to divert; further, if the trip took longer, then drivers were more likely to divert (*5–7*).

A sign for perception of congestion on the alternate route is negative, suggesting that worse congestion on the alternate route inhibits drivers from taking it. This finding underscores the importance of information about congestion on alternate routes. The sign for the number of alternates used is positive, which implies that more knowledge of alternate routes encourages drivers to divert. The number of alternatives used indicates the variety of paths connecting home and work. This variable is an indicator both of cognitive maps of drivers (i.e., driver familiarity with alternate routes) and of the perceived alternatives for diversion. It is inferred that richness of cognitive maps influences diversion behavior.

Among driver characteristics, gender and self-assessment statements about risk behavior are significant. Females were less likely to divert than were males (*6*). The adventure and discovery aspects of driver personality, intended to capture the willingness to take risks and an interest in discovery and exploration, are positively associated with diversion. This factor represents the sum of scores normalized by the number of statements, and the scores vary between 0 and 4. A score of 0 indicates no interest in adventure and discovery, and a 4 indicates great enthusiasm for adventure and discovery.

The stated preference index (SPI) was created by summing and normalizing the responses to the stated preference questions. The scores on this index vary between 0 and 4. Zero indicates that the respondent would definitely take the usual route in all scenarios investigated, and 4 indicates that the respondent would definitely take the alternate route in all scenarios. This index is a measure of a driver's propensity to divert. The SPI has a positive sign, as expected. It is inferred that stated preferences may be good representations of revealed preferences.

SUMMARY AND CONCLUSIONS

Commuter response to delay was investigated in real-life situations to explore the effects of factors such as driver and trip characteristics, route attributes, traffic information, and environmental conditions on driver response to delay. The main advantage of this research design is that it supports investigating driver response to delay and econometric modeling of diversion behavior at the disaggregate level.

En route diversion behavior was found to be influenced by several factors, including source of traffic information, length of delay, gender, travel time, number of alternate routes used, congestion on the alternate route, residential location (city or suburbs), self-evaluation statements about risk behavior (personality), and stated preferences about diverting.

The key finding is that real-time traffic information influences en route diversion behavior. Drivers were more likely to divert when they received delay information through radio reports than when they observed the delay. Thus, real-time traffic information provides a basis for making en route diversion decisions, and drivers actually shift their routes in response to radio traffic information. The use of public and private resources devoted to the collection and dissemination of real-time traffic information can produce benefits for drivers. Other findings are summarized as follows:

- The relationship between expected length of delay and diversion to alternate routes derived from the revealed preference method does not show the clean positive relationship found in stated preference studies, possibly because the true magnitude of a delay is often not obvious to a driver in advance.
- Most drivers who diverted believed that they saved travel time by diverting. The self-reported saving in travel time for drivers who diverted in response to delay averaged about 10 min. Although the self-reported savings may be rationalizations of drivers' diversion decisions, drivers may need some minimum travel time saving, here about 10 min, to justify diverting to alternate routes.
- The number of alternate routes known to drivers, that is, cognitive maps, influence their response to delay.
- A driver's inclination toward adventure and discovery encouraged en route diversion.

If society is to get the most congestion relief from real-time traffic management, transportation planners need to understand how drivers make en route diversion decisions—starting with how drivers use currently available information. This study has shown that it is feasible to conduct empirical studies of such behavior. The results indicate that real-time information is influential in diversion decisions. Moreover, length of delay and perception of traffic congestion on the alternate route influence en route diversion decisions. Therefore, it is recommended that, in the short term, improvements in real-time traffic information focus on developing and disseminating predictions of delay duration due to incidents and the congestion levels on the alternate routes surrounding the incident. This action would require monitoring traffic conditions on the alternate routes along major roadways. Providing clearer information on delays and congestion may contribute favorably to congestion-reducing demand management strategies. In the Chicago area (at least), the provision of real-time traffic information may offer significant untapped potential for better serving suburb-to-CBD commuters. This objective might be achieved by extending the system of detectors further into the suburbs, as well as by monitoring major arterial streets.

ACKNOWLEDGMENTS

Financial support and computer resources were provided by The Transportation Center, Northwestern University, Evanston, Illinois. The authors thank the anonymous referees for their constructive criticism.

REFERENCES

1. J. McDermott. Freeway Surveillance and Control in Chicago Area. *Transportation Engineering Journal*, TE 3, May 1980, pp. 333–348.
2. Working Groups of Mobility 2000. *Reports on Major Aspects of IVHS*. Texas Transportation Institute, Communications Division, Texas A&M University, College Station, 1990.
3. C. Dudek, C. Messer, and H. Jones. Study of Design Considerations for Real-Time Freeway Information Systems. In *Highway Research Record 363*, HRB, National Research Council, Washington, D.C., 1971.
4. R. Huchingson and C. Dudek. Delay, Time Saved, and Travel Time Information for Freeway Traffic Management. In *Transportation Research Record 722*, TRB, National Research Council, Washington, D.C., 1979, pp. 36–40.
5. Y. Stephanedes, E. Kwon, and P. Michalopoulos. Demand Diversion for Vehicle Guidance, Simulation, and Control in Freeway Corridors. In *Transportation Research Record 1220*, TRB, National Research Council, Washington, D.C., 1989, pp. 12–20.
6. F. Mannering. Poisson Analysis of Commuter Flexibility in Changing Routes and Departure Times. *Transportation Research*, Vol. 23B, No. 1, 1989, pp. 53–60.
7. H. Mahmassani, C. Caplice, and C. M. Walton. Characteristics of Urban Commuter Behavior: Switching Propensity and Use of Information. *Transportation Research Record 1285*, TRB, National Research Council, Washington, D.C., 1990, pp. 57–69.
8. K. Heathington, R. Worrall, and G. Hoff. Attitudes and Behavior of Drivers Regarding Route Diversion. In *Highway Research Record 363*, HRB, National Research Council, Washington, D.C., 1971, pp. 18–26.
9. R. Huchingson, J. Whaley, and N. Huddleston. Delay Messages and Delay Tolerance at Houston Work Zones. In *Transportation Research Record 957*, TRB, National Research Council, Washington, D.C., 1984, pp. 19–21.
10. M. Haselkorn, J. Spyridakis, L. Conquest, and W. Barfield. Surveying Commuter Behavior as a Basis for Designing Motorist Information Systems. Presented at 1st Vehicle Navigation and Information Systems Conference, IEEE, Toronto, Canada, 1989.
11. R. Huchingson, R. McNees, and C. Dudek. Survey of Motorist Route-Selection Criteria. In *Transportation Research Record 643*, TRB, National Research Council, Washington, D.C., 1977, pp. 45–48.
12. E. Daniels, M. Levin, and J. McDermott. Improving Commercial Radio Traffic Reports in the Chicago Area. In *Transportation Research Record 600*, TRB, National Research Council, Washington, D.C., 1976, pp. 52–57.
13. E. Shirazi, S. Anderson, and J. Stesney. Commuter's Attitudes Toward Traffic Information Systems and Route Diversion. In *Transportation Research Record 1168*, TRB, National Research Council, Washington, D.C., 1988.
14. C. Dudek, G. Weaver, D. Hatcher, and S. Richards. Field Evaluation of Messages for Real-Time Diversion of Freeway Traffic for Special Events. In *Transportation Research Record 682*, TRB, National Research Council, Washington, D.C., 1978.
15. C. Dudek and R. Huchingson. *Manual on Real-Time Motorist Information Displays*. Report FHWA-IP-86-16. FHWA, U.S. Department of Transportation, 1986.
16. C. Dudek, R. Huchingson, and R. Brackett. Studies of Highway Advisory Radio Messages for Route Diversion. In *Transportation Research Record 904*, TRB, National Research Council, Washington, D.C., 1983.
17. C. Dudek and R. Huchingson. *Human Factors Design of Dynamic Displays, Volume 1*. Report FHWA/RD-81/039. FHWA, U.S. Department of Transportation, 1982.

18. S. Richards, W. Stockton, and C. Dudek. Analysis of Driver Responses to Point Diversion for Special Events. In *Transportation Research Record 682*, TRB, National Research Council, Washington, D.C., 1978, pp. 46–52.
19. C. Dudek, W. Stockton, and D. Hatcher. Real-Time Freeway-to-Freeway Diversion: The San Antonio Experience. In *Transportation Research Record 841*, TRB, National Research Council, Washington, D.C., 1982, pp. 1–14.
20. D. Roper, R. Zimowski, and A. Iwamasa. Diversion of Freeway Traffic in Los Angeles: It Worked. In *Transportation Research Record 957*, TRB, National Research Council, Washington, D.C., 1984, pp. 1–4.
21. J. M. Turner, C. Dudek, and J. Carvell. Real-Time Diversion of Freeway Traffic During Maintenance Operations. In *Transportation Research Record 683*, TRB, National Research Council, Washington, D.C., 1978, pp. 8–10.
22. R. Golledge and R. Stimson. *Analytical Behavioral Geography*. Croom Helm, New York, 1987.
23. S. Freundschuh. Does "Anybody" Really Want (or Need) Vehicle Navigation Aids? Presented at First Vehicle Navigation and Information Systems Conference, IEEE, Toronto, Canada, 1989.
24. *A Survey of Persons Who Drive and Park in the Chicago Central Business District*. Chicago Area Transportation Study, Chicago, Ill., 1987.
25. A. Khattak. *Factors Influencing Driver Response to Incidents*. Ph.D. dissertation. Civil Engineering Department, Northwestern University, Evanston, Ill., 1991.
26. G. Ergun. *The Effects of Personality on Recreational Travel Behavior*. Ph.D. dissertation. Civil Engineering Department, Northwestern University, Evanston, Ill., 1979.
27. J. Horowitz, F. Koppelman, and S. Lerman. *A Self Instructing Course in Disaggregate Mode Choice Modeling*. Report IA-11-0006. UMTA, U.S. Department of Transportation, 1986.
28. M. Wachs. Relationships Between Drivers' Attitudes Toward Alternate Routes and Driver and Route Characteristics. In *Transportation Research Record 197*, HRB, National Research Council, Washington, D.C., 1967, pp. 70–87.
29. M. Vaziri and T. Lam. Perceived Factors Affecting Driver Route Decisions. *Journal of Transportation Engineering*, Vol. 109, No. 2, March 1983.
30. L. Haefner and L. Dickinson. Preliminary Analysis of Disaggregate Modeling in Route Diversion. In *Transportation Research Record 527*, TRB, National Research Council, Washington, D.C., 1974, pp. 66–72.

DISCUSSION

R. D. HUCHINGSON
Industrial Engineering Department, Texas A&M University, College Station, Tex. 77843.

The authors have made a significant contribution to the state of the art in commuter responses to delay. However, I would like to comment on several issues that the authors have raised.

The authors make a distinction between the stated performance approach and the revealed preference approach. The former is defined as what respondents report when asked if they would divert in response to a particular delay. This is an estimate of probable behavior. The latter is based on respondent reports of driver behavior in real-life situations when asked about their actions in previously encountered situations. This is a report of actual behavior in a recalled situation.

Whether or not a driver did or did not divert in a recently recalled situation depends heavily on the information provided the driver by a communication system regarding the extent of delay and the severity of the problem. Real-time changeable message sign (CMS) information from a traffic control center can be very explicit regarding the delay the driver could expect. However, listening to commercial radio reports or simply observing the traffic ahead may leave unspecified exactly the delay that will be encountered. In the first instance, the information may be assumed to be reliable. In the latter instance, it is only a gross estimate.

Research reported by Huchingson and Dudek (*1*) was based on a multistate survey of responses regarding whether or not respondents as drivers would divert to various specified intervals of delay. The objective was to determine those intervals of delay that should be displayed on a CMS. The relationship between delay intervals and reported diversion was found to be S-shaped: few were willing to divert to short delays and virtually everyone was willing to divert to lengthy delays. The curves varied somewhat with the message, i.e., Delay or Time Saved in taking an alternate route.

Two of the more astounding findings of the reviewed study are (a) no more than 50 percent of drivers stated that they would divert regardless of the length of delay; and (b) at durations in excess of 40 min actually fewer drivers would divert than at 31 to 40 min of delay (see Figure 3).

The authors account for this seemingly irrational behavior by stating that in real-life situations drivers cannot know with certainty the length of delay in advance, and, hence, "may not display the same level of rationality to the researcher." They also note that sometimes delay is perceived incrementally and that there were only 24 observations in the delay category of 40 min or more.

The study and the one by Huchingson and Dudek (*1*) may have been based on different assumptions. The S-shaped finding was based on the assumption that reliable information did exist and that given this information drivers were to report how they would react in terms of diversion. In the present study, accurate information was lacking.

Although the findings are not contradictory, there is a risk that the casual reader may conclude that 50 percent of drivers would not divert in the real-world situation and that agencies cannot expect much more voluntary diversion than this. Huchingson and Dudek (*1*) found that in the multistate study only about 4 percent would not divert given that they had reliable information that the delay was 1 hr or more. In the reviewed study, the drivers were not reacting to delay information but rather reporting their diversion behavior in a situation where accurate information was absent.

The data regarding perceived loss or savings in travel time from either diverting or not diverting should be interpreted also in terms of the information available to the drivers in the sample. It was concluded from the study that the average saving in travel time for those who diverted is 9.6 min and that the expected loss in travel time for those who did not divert is 4.2 min. These values were deemed reasonable by the authors because other researchers have also reported that the median delay before diversion varies between 5 and 27 min.

The authors correctly note that the reported estimates may well be exaggerated by drivers to justify or rationalize their decisions. It is true that the drivers had little if any reliable feedback from the information system regarding how much time they actually saved or lost in the situations posed. The specific responses given were based largely on subjective factors unrelated to the actual savings or losses of time.

The authors also conducted a stated preference study with

12 posed situations and reported the percentage of drivers who would take the usual route and diversion route. Nine of the scenarios dealt with 15-min delays. Ten of the best alternate route conditions were given as "no delay"; the other two pose "no information."

The last three scenarios pose 10, 15, and 20 min delay. The percentage data from these delays are plotted in Figure 4 to show how drivers stated behavior differed from their "revealed behavior." This comparison may have been an afterthought for, if the intent was to provide a comparison, the scenarios should have included posed delays of 25, 35, and 50 min. As a control group, it is incomplete.

The upper curve percentages are similar to those found by Huchingson and Dudek (*1*) for 10, 15, and 20 min of delay (18, 34, and 63 percent diversions, respectively).

Huchingson and Dudek (*1*) found 83 percent diversion to 30 min of delay and 95 percent diversion to 60 min. Only 1 percent more (96 percent) would divert to 120 min. Four percent refused to divert. It would seem unreasonable for a large percentage of drivers to sit in stalled traffic for 1 hr given they knew of alternate routes with traffic moving.

The authors pose an interesting theory regarding how anticipated delay relates to knowledge of additional delay. This "proportional delay theory" is illustrated by the following cited example. "For example, a 5 minute increase in a 10 minute delay (a 50% increase in length of delay) has the same effect on a driver's diversion decision as adding 15 minutes to a 30 minute delay." In other words, the longer one expects to be delayed in traffic, the greater the percent increment in delay required before the driver will be concerned. The authors state that this effect applies up to 1 hr of delay. Thus, a logarithmic rather than a linear transform was used in the model.

Although this theory has some plausibility (generalizing from Weber's Law in psychophysics), no data were cited to merit the adoption of the model other than "the log transform gave slightly better statistical results." This theory should be tested by a parametric study presenting various levels of delay and noting if there was reduced time sensitivity as units of delay and trip length increase.

There are interesting findings from the study related to driver characteristics (gender and personality) and the length of the commute. However, the traffic engineer attempting to apply the findings to real-world problems must be fully aware that the drivers in the study had limited information on the actual delay times on the usual and alternate routes. Further, the study findings may be situationally specific. Characteristics of the alternate routes in the Chicago area in terms of their desirability for diversion may be quite different from those in another major metropolitan freeway system. There may also be seasonal variations (e.g., heavy snowfall affecting all facilities). Caution must be applied when attempting to generalize from the data to other cities. The findings should be accepted recognizing that the drivers had precious little real-time information on time saved or lost or feedback in the time domain. The numbers reported in the survey should be viewed in that light.

REFERENCE

1. C. Dudek and R. Huchingson. Delay, Time Saved, and Travel Time Information for Freeway Traffic Management. In *Transportation Research Record 722*, TRB, National Research Council, Washington, D.C., 1979, pp. 36–39.

Publication of this paper sponsored by Committee on User Information Systems.

Legibility and Contrast Requirements of Variable-Message Signs

MICHÈLE COLOMB AND ROGER HUBERT

New technologies such as optic fibers and light-emitting diodes are now used for information matrix signs. A field study was carried out to evaluate the best conditions for the legibility of these signs during the day and at night. For legibility criteria, the contrast between the letters and the sign background is chosen for daylight conditions and the luminance of the letters for night conditions. The performance of some commercially available signs is compared with the study results.

The variable-message sign (VMS) is being used increasingly on main roads and motorways. On those signs, optical fibers are not being followed by light-emitting diodes. The special characteristics of these products call for a revision of the French specifications to complete the recommendations on their luminous intensities and colors. The focus here is on the photometric aspects of signs delivering alphanumeric messages made up of 5- by 7-point dot matrix characters. An initial investigation was carried out by simulation on video monitors (*1*), but the luminosity limits of the monitors were such that only night visibility conditions could be simulated. Hence, an experimental study has been carried out on a prototype sign to determine, in a more realistic manner, the conditions of optimum visibility both in daylight and at night.

REVIEW OF LITERATURE ON VMS LEGIBILITY

There is a tendency to characterize a fiber optic or diode VMS consisting of luminous points by the luminous intensity emitted by each point of the matrix. But, if such signs are observed from a long distance (100 to 200 m), the characters exhibit a continuous appearance under most traffic conditions.

Which photometric property should then be chosen to characterize the legibility of such signs? Luminous intensity is closest to technological reality, but luminance and contrast are more closely correlated with human vision and with the criteria of legibility applied to other signs (*2*).

The latter two factors were chosen. Kerr et al. (*3*) have carried out an experiment on the legibility of VMS through laboratory simulation. By analyzing the observers' response times, they found that reading performance was best with a contrast of approximately 7.

Moreover, reading is generally dependent on the visual acuity of the observers. This factor varies substantially with the luminance of the character displayed and with the contrast between the character and its background (*2*). Van Meeteren et al. (*4*) have found that a luminance ratio of approximately 10 between the letter and the background provides optimum visual acuity.

The results of this investigation, obtained in daylight conditions, are presented in terms of contrast C', equal to luminance ratio L/Lf, where L is the luminance of the letter and Lf is the luminance of the background. This relation takes ambient luminosity into account in the luminance of the background of the sign.

Other authors, among them Padmos et al. (*5*), have chosen the luminance of the characters as the VMS legibility criterion, stated as a function of the luminance of the horizon, both in daylight and at night. This factor also takes ambient luminosity into account. On the basis of this investigation, carried out at an actual site at an observation distance of 100 m, two levels are recommended for day and night conditions: a luminance of 4000 cd/m^2 in daylight and 100 cd/m^2 at night. At night, the luminance of the backgrounds, of the VMS tends toward zero, so the luminance of the character was chosen as the criterion, as used for other types of sign.

For example, for comfortable reading of illuminated signs, Allen et al. (*6*) recommend a luminance of 30 to 300 cd/m^2 in rural areas (where there is little or no illumination) and 300 to 1500 cd/m^2 in highly illuminated urban areas. For retroreflecting signs, Woltman and Szczech (*7*) report an optimum character luminance of 75 cd/m^2 to ensure good legibility at night. Reading becomes just possible at only a few cd/m^2.

The previous study of the legibility of dot matrix characters carried out by simulation on video monitors (*1*) indicated more than 70 percent responses with a luminance of between 30 and 230 cd/m^2. Therefore, the criterion of character luminance is used to provide the night results.

EXPERIMENTAL INVESTIGATION

Description of Sign

The experimental investigation used a prototype sign built for that purpose by the optronics laboratory of the University of Poitiers. It has a single block of diodes, 320 mm high, on which any of the 26 letters of the alphabet can be displayed (see Figure 1). Each letter is defined by a 5- by 7-point array.

Each point of the matrix may consist of a variable number of diodes: 1, 4, 9, 16, 25, or 36. The luminous intensity of each point of the matrix varies according to the number of diodes lit and according to the electric power delivered to each diode. In practice, the luminous intensity was varied

Laboratoire Central des Ponts et Chaussées, 58 boulevard Lefebvre, 75015 Paris, France.

FIGURE 1 Experimental sign displaying the letter E, with each point of the matrix consisting of nine diodes.

from approximately 0.02 to 1.5 cd per point under the night observation conditions and from 0.2 to 8 cd per point under the daylight observation conditions.

During the tests, the vertical illumination received by the sign and the background luminance of the sign when off were measured. On the different test days, the vertical illumination ranged from 5000 to 60 000 lx and the background luminance from 100 to 500 cd/m^2.

Experimental Procedures

The sign was shown to 27 observers, in groups of 3 from a stopped vehicle 200 m away. The 26 letters of the alphabet were presented in random order at six different matrix point sizes and six different levels of luminous intensity. Altogether, 180 configurations were displayed, each for 2 sec. After each presentation, the observers recorded the letter read on a form. The same procedure was followed both for daylight and for night conditions. The observers consisted of 8 women and 19 men, between 22 and 65 years old. The drivers' vision was checked by measuring visual acuity and sensitivity to contrast.

Daylight Results

The answer forms were analyzed by counting correct answers, incorrect answers, and failures to answer for all individuals and each luminous level and size of matrix point. As a result of the thinking about readability criteria discussed previously, the daylight results were reported by the percentage of correct answers versus contrast C' (see Figure 2), so that they could be compared with the results given in the literature.

In daylight the true luminance (L) of the character is the sum of the internal luminance (Li) of the sign and the external luminance (Le) resulting from ambient illumination, which is equal to the background luminance (Lf). Internal luminance (Li) is calculated from the luminous intensity measured in the photometrics laboratory by the following equation:

$$\text{Li} = (\text{Ip} \times 35)/S \qquad (1)$$

where Ip is the luminous intensity per point of the matrix and S is the area of the block of diodes containing the 35 points. Thus,

$$C' = L/\text{Lf} = (\text{Li} + \text{Lf})/\text{Lf} \qquad (2)$$

Because the background luminance was not measured continually in the course of the tests and varied with the ambient illumination, the average value of Lf = 200 cd/m^2 was used for the contrast calculation. The true luminance of the letters ranged from approximately 280 to 4090 cd/m^2.

The various symbols of Figure 2 represent the six sizes of the points of the matrix.

The experiment did not reveal any substantial influence of point size on reading performance, whereas the influence of contrast C' can be seen in the rapid increase in correct answers: from 10 to 50 percent as C' increases from 1.5 to approximately 3. The percentage of correct answers continues to rise with increasing contrast, leveling off at about 85 percent for a contrast between 8 and 20. No values were measured beyond 20. The corresponding luminance (L) of the character is between 1500 and 4000 cd/m^2. These values are perfectly compatible with the results of Padmos et al. (*5*).

Night Results

In accordance with the thinking about night legibility criteria, the results are given as the percentage of correct answers by the observers versus the luminance of the character. At night there is little illumination of the experimental site and no oncoming vehicles. The results were obtained by the same procedures as for daylight. The various symbols on Figure 3 represent the different matrix point sizes. At night, as in daylight, the experiment did not reveal any significant influence of this factor. During the night experiments, the luminance ranged from 9 to 730 cd/m^2. Overall performance was 60 percent, with a large dispersion, and revealed no significant change in the percentage of correct answers with increasing luminance. This performance level, lower than in daylight, is probably explained by the observers' loss of visual acuity at night.

Most of the observers judged the highest luminance levels to be uncomfortable, but this perceived discomfort did not affect reading performance, probably because each letter was presented long enough for the individuals' vision to adapt to these slightly more difficult reading conditions.

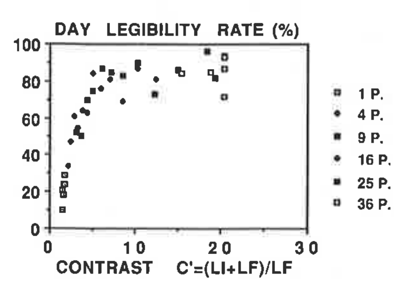


FIGURE 2 Percentage of letters correctly read in daylight versus contrast.

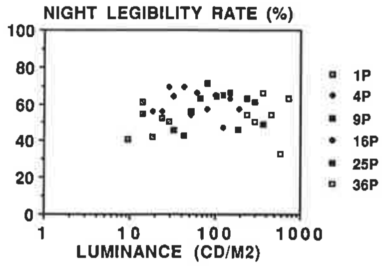


FIGURE 3 Percentage of letters correctly read at night versus luminance.

The night results do not allow precise values of luminance required for reading to be determined. The previous study by Mazoyer and Colomb (*1*), in which a simulation was used, indicated a narrower range of luminances ($30 < L < 230$ cd/m^2) that are perfectly compatible with the other data examined.

PHOTOMETRIC PERFORMANCE OF THE VMS

In this section the results of this experiment and the results reported in the literature are compared with the actual performance of the products.

Measurement Method

A testing method to evaluate the luminous performance of products subject to approval was developed in the photometric laboratory. The measurement includes a number of steps.

First, a prototype sign capable of displaying three characters is placed on a rotating table. A photometric cell that can move on a column is used to measure the luminous intensity of the characters and, hence, to deduce the luminous intensity per matrix point (see Figure 4). The measurement is automated; it begins on the axis perpendicular to the plane of the sign and is continued to ± 16 degrees horizontally. The data (angle, luminous intensity) are transmitted to the microcomputer-controlled acquisition system, and the results are printed out immediately as tables or graphs.

Second, the background luminance of the sign is measured (see Figure 5). The sign is once again placed on the rotating table. It is kept off and is illuminated by a light source simulating the sun, from a direction 20 degrees above the normal to the sign in the vertical plane. The measurement of luminance L is made in the axis of the sign by a luminancemeter at a given vertical illumination (EV). Unfavorable illuminance conditions are defined by EV = 80,000 lux. The corresponding background luminance is calculated using the following relationship:

$$\text{Lf} = (L \times 80\,000)/\text{EV} \qquad (3)$$

Finally, the color of the luminous message is measured on the illuminated panel using a spectrophotometer in the axis perpendicular to the front of the sign.

Performance in Daylight

By setting a minimum contrast threshold and determining the background luminance of the sign, it is possible to calculate the luminance that a character should have to be readable and, from that value, to deduce the luminous intensity per matrix point for a given character height.

Assume a contrast threshold of 3, corresponding to 50 percent correct answers in the investigation. Under unfavorable conditions of illumination, with a vertical illumination of 80 000 lux, the background luminance of signs can easily be as much as 1500 cd/m^2 with a diffusing front surface. Under these conditions the luminance of the character must exceed 4500 cd/m^2 and the internal luminance 3000 cd/m^2. Assuming a character height of 400 mm, the corresponding luminous intensity per matrix point must be greater than 10 cd.

The luminous intensity values vary with the technology. The values per matrix point Ip as measured in the photo-

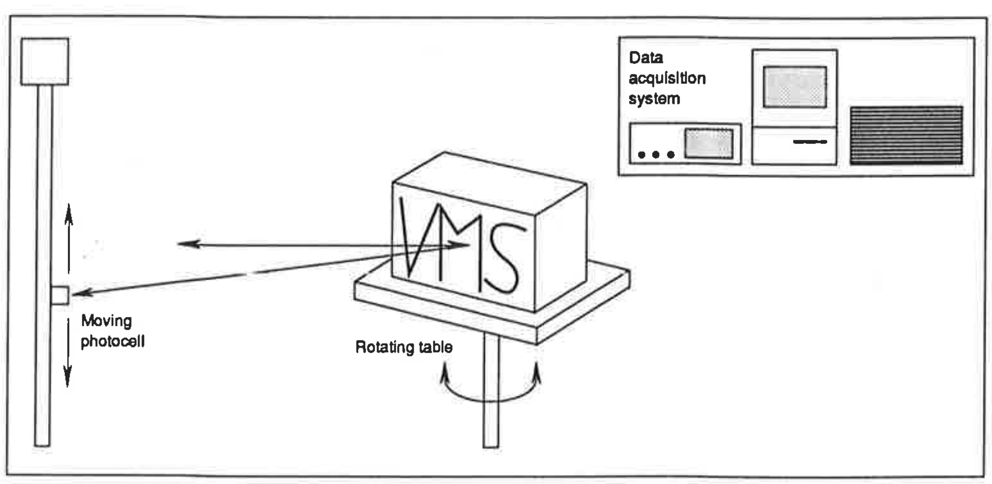


FIGURE 4 Measurement of luminous intensity.

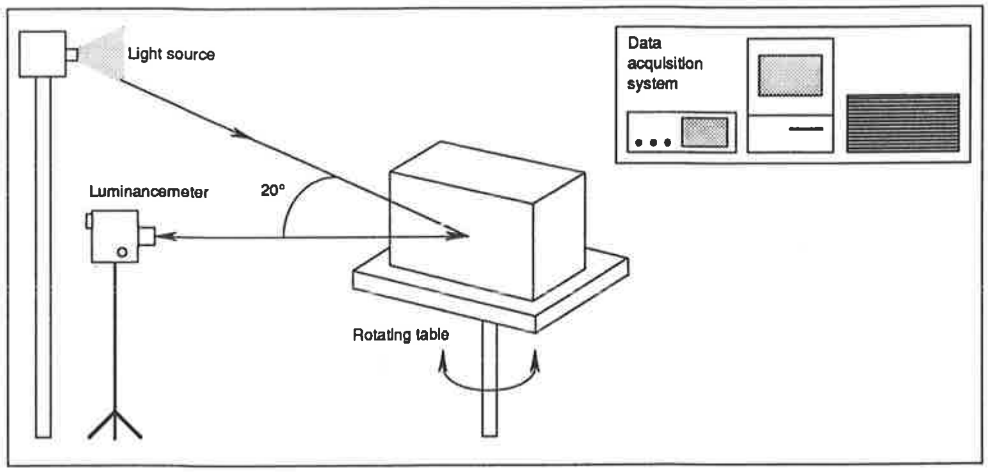


FIGURE 5 Measurement of luminance of front of sign.

metrics laboratory in the direction perpendicular to the front of the sign are as follows:

- For fiber optic signs, Ip ranges from 10 to 40 cd according to the power of the source, the number of fibers per matrix point, and the color of the filter used (yellow in some cases).
- For diode signs, Ip is a few candela and varies with the number of diodes per point and their color (red diodes are more powerful than the others).

A value of 3.6 cd was measured on a sign for which each matrix point consisted of 8 red diodes and 8 green diodes. A value of 4.3 cd was measured on a sign for which each matrix point consisted of 20 yellow diodes and 20 green diodes.

The luminous intensity of the diodes decreases with increasing temperature; at constant ambient temperature, it decreases after the diodes are switched on because of heating. The values given are for the standard measurement conditions, with a temperature of 20°C and measurement made 20 min after lighting.

Judged by the design assumptions stated, only the fiber optic technology provides adequate legibility, with Ip greater than 10 cd.

To consider another example, the initial assumption concerning contrast (contrast threshold equal to 3) is retained, but assumptions concerning the front of the sign are changed. Consider a sign having a front that diffuses very little, measured under the same conditions of illumination or with weaker illumination. The background luminance is only about 500 cd/m^2. If a character height of 400 mm is assumed, the luminous intensity per point must exceed about 3 cd. In this example, fiber optic and light emitting diodes are both satisfactory.

Thus, when the siting of a VMS is being considered, the natural conditions of illumination to which the sign may be exposed (unfavorable conditions) must be examined along with the photometric properties of the sign (i.e., its luminous intensity), in conjunction with the optical properties of the front of the sign.

Performance at Night

From the luminance ranges given, it is possible to calculate the corresponding luminous intensity for characters of a given height. If, for example, the two luminance ranges specified by Allen (*6*) are used for the two ambient luminosities, with a character height of 400 mm, the luminous intensity per dot, at night, must be between 0.1 and 1 cd in a zone of little or no illumination and between 1 and 5 cd in a highly illuminated urban zone. Most of the signs available on the market provide a luminous intensity that is greater than the required values. Indeed, care should be taken to reduce these intensities for night operation.

CONCLUSIONS

The criteria of contrast and luminance can be used to specify the conditions that a VMS must satisfy to be readable in daylight and at night. A review of the literature revealed overall agreement among the ranges of values obtained in the various investigations.

In practice there may be a problem of legibility in daylight with some VMSs. An overall photometric analysis of the sign can help correct this problem by considering not just the emitted luminous intensity but also the luminance of the front of the sign.

The photometric problems posed by this new type of sign have been reviewed. It now remains to settle the question of color raised by the use of diodes, which display spots of color outside the ranges traditionally accepted by the International Commission on Illumination for other luminous signaling equipment.

ACKNOWLEDGMENTS

The authors would like to thank Mrs. Doré-Picard of the Institut National de Recherche sur les Transports et leur Sé-

Distribution of Driver Spare Glance Durations

George T. Taoka

The statistical distributions of driver off-roadway spare glance durations for six in-vehicle and one out-of-vehicle viewing tasks are generated. An analytical model is fitted to experimental data from several sources to yield these distributions. Of the seven types of driver viewing tasks, two are required for navigation purposes, whereas five need only be viewed occasionally and usually under light traffic conditions. The two navigational tasks are viewing the rearview and left-side mirrors. The other five tasks include viewing the speedometer, temperature, and defroster gauges, glancing at the radio, and reading roadway name signs outside the moving vehicle. Mean, standard deviation, and median values are given for these viewing task durations, and probability distribution values are tabulated. A goodness-of-fit test comparing model values to experimental data is also conducted for the driver task of viewing the car radio.

The study of driver glance durations in scanning the roadway scene ahead of the moving vehicle has been a popular topic of human factors research, and much is known about this important topic. In contrast, comparatively little research has been conducted on driver spare glance durations for off-roadway viewing tasks, such as viewing dashboard displays and rearview mirrors. One of the earliest studies was published by Mourant and Rockwell (*1*). This study focused on the differences in the scanning behavior of novice and experienced drivers in viewing mirrors and the speedometer while driving under freeway and neighborhood conditions. Mourant and Rockwell noted that experienced drivers used the rearview mirror more than the novice drivers did but that the reverse was true for the use of the speedometer. They also concluded that the spare search and scan patterns of most novice drivers were generally inadequate for safe driving on public highways. Mourant and Donohue (*2*) later focused on the drivers' glance durations to the rearview and left-side mirrors while executing both left- and right-lane changing and merging maneuvers.

Within the past 5 years, several studies have been published that include statistical parameters of driver spare glance durations for several important viewing tasks. The availability of these statistical parameter estimates allows the distribution of these glance durations to be computed on the basis of an appropriate analytical model. Statistical distributions of driver spare glance durations to viewing two mirrors, four dashboard displays, and one out-of-vehicle target are presented. A goodness-of-fit test is applied between model estimates and experimental data for one of these viewing tasks.

Department of Civil Engineering, University of Hawaii at Manoa, 2540 Dole Street, Honolulu, Hawaii 96822.

EXPERIMENTAL STUDIES OF SPARE VIEWING TIMES

Nagata and Kuriyama (*3*) published viewing times of automobile drivers using side mirrors. They found the mean glance durations of drivers to be 0.69 sec for the near-side mirror and 1.38 sec for the far-side mirror. The difference in viewing times was found because the near-side mirror was at a 41.5-degree angle from the longitudinal axis, whereas the far-side mirror was at a 70-degree angle from the longitudinal axis parallel to the roadway. Significant differences existed in the glance times of the three groups, which included driving instructors, male students under 20 years of age, and female students over 30 years of age. The driving instructors exhibited significantly shorter glance times than the other two groups did.

In 1987, two extensive experimental studies were presented at the Second International Conference on Vision in Vehicles held in Nottingham, England. Wierwille et al. (*4*) presented data on glance time durations of drivers viewing 26 different tasks. A total of 32 drivers drove a specially instrumented 1985 Cadillac Deville on public roadways in Virginia, under light and moderate traffic conditions. The drivers ranged in age from 18 to 73 and varied in driving experience from 2,000 to 40,000 mi of driving per year. The average glance times exhibited by this group ranged from 0.62 sec in viewing the speedometer to 1.63 sec in reading a roadway name sign as the vehicle approached and moved past the sign. Wierwille et al. concluded that drivers over 50 years of age required longer glance times and made greater errors in reading dashboard displays than did the younger subjects. Some conventional tasks, such as turning on and viewing the radio, required more time than viewing the speedometer. In general, viewing tasks requiring greater complexity in interpretation required longer glance times than those requiring only simple identification.

The second study, presented by Rockwell (*5*), contained data on driver glance durations in viewing the radio and the left-side mirror. A total of 106 drivers were tested over a 6-year period on highways and expressways under light to moderate traffic conditions. The subjects were equally divided according to gender and age to represent a composite of the general driving population. Average glance times were shortened by an average of about 20 percent under higher traffic density conditions than under lower, less crowded conditions.

curité for the ergonomic aspects of the study; Messrs. Canestrelli, Richard, and Le Fur of the Laboratoire Central des Ponts et Chaussées (LCPC) for the technical organization of the experiment; and Messrs. Carta and Peybernard of LCPC for data processing. They also thank their colleagues at LCPC who kindly took part in the study as observers.

REFERENCES

1. M. Mazoyer and M. Colomb. Investigation of Legibility of Dot-Matrix Signs by Simulation on a Video Screen. *Proc., 21st Session, International Commission on Illumination*, Venice, June 1987.
2. *Roads Signs*. Publication 74. International Commission on Illumination, 1988.
3. J. Kerr, R. Snelgar, T. Jordan, P. Emmerson, and P. Linfield. Optimum Display Factors for Light-Emitting Variable-Message Road Signs. *Proc., V.I.V.II*, Sept. 1987, pp. 277–288.
4. A. Van Meeteren, H. Leebeek, and N. H. Blokland de Graaf. *Legibility of Internally Illuminated Highway Signs*. Report IZF 1968, C2. The Netherlands Organization for Applied Scientific Research TNO, 1968.
5. P. Padmos, T. Van den Brink, J. Alferdinck, and E. Folles. Matrix Signs for Motorways: System Design and Optimum Light Intensity. *Proc., V.I.V.II*, Sept. 1987, pp. 269–276.
6. T. M. Allen et al. Luminance Requirements for Illuminated Signs. *Highway Research Record 179*, HRB, National Research Council, Washington, D.C., 1967, pp. 16–37.
7. H. L. Woltman and T. J. Szczech. Sign Luminance as a Methodology for Matching Driver Needs, Roadway Variables, and Signing Materials. In *Transportation Research Record 1213*, TRB, National Research Council, Washington, D.C., 1989.

Publication of this paper sponsored by Committee on User Information Systems.

It therefore appears that driver glance durations are a function of vehicular roadway density.

DISTRIBUTIONS OF SPARE GLANCE DURATIONS

The statistical parameters of radio viewing glance times were presented as 1.27, 1.44, and 0.5 sec, for the median, mean, and standard deviation values respectively, of 1,250 recorded glances. This finding yields a value of 0.376 for the dispersion parameter associated with the lognormal probability density model used here. The cumulative probability distribution estimates ranged from 0.68 sec for the 5th percentile to 2.35 sec at the 95th percentile.

In this study, women took shorter glances to the radio than men did. Older drivers also recorded longer times than those of their younger counterparts. Radio glance times were longer for viewing the radio than for viewing other dashboard displays, because radio operation required visual discrimination, whereas viewing other displays only required visual detection—a shorter process. Although station selection required an average of 1.50 sec, tuning the volume required only an average time of 0.97 sec. The degree of visual discrimination was, therefore, an important factor in determining the length of the driver glance duration required for a specific visual task.

The average time spent glancing at the left-side mirror was 1.10 sec, which is considerably shorter than the average radio glance duration. The median glance duration was 1.06 sec, with a standard deviation value of 0.3 sec. The corresponding dispersion parameter value was 0.268 sec. The 5th-percentile value of 0.68 sec was identical with the value for radio glances, but the 95th-percentile value of 1.65 sec was significantly shorter than the corresponding radio glance value. The older drivers, however, took shorter glances than their younger counterparts when viewing the left-side mirror. No significant differences in gender were noted for this viewing task.

The glance times presented for viewing the rearview mirror were significantly shorter than those required for viewing the left-side mirror. These values were 0.75 sec for the average, 0.68 sec for the median, and 0.36 sec for the standard deviation. The 5th- and 95th-percentile values were 0.32 and 1.43 sec, respectively.

The mean, median, and standard deviation parameters in viewing the speedometer were 0.62, 0.49, and 0.48 sec, respectively, with a dispersion estimate of 0.685.

The corresponding statistical parameter estimates for viewing the temperature gauge, the defroster gauge, and roadway name signs are presented in Table 1. The average values range from 1.10 sec for the temperature gauge to 1.63 sec for reading roadway name signs. The task of viewing roadway name signs was chosen to represent at least one task involving viewing a target outside the moving automobile. This task required visual discrimination as well as visual detection, thereby lengthening the time required to complete it.

LOGNORMAL PROBABILITY DENSITY MODEL

The probability density model of driver glance durations, $f(t)$, is lognormal distributed if the standard normal variable (z) defined by the equation

$$z = \frac{\ln(t) - \ln(t_m)}{d} \tag{1}$$

is normally distributed with a mean value of 0.0 and a standard deviation value of 1.00. In this equation t_m is the median, or 50th percentile, value of the driver glance time (t) in seconds. The dispersion parameter (d) satisfies the following relationship:

$$d^2 = \ln\left[1 + \left(\frac{s}{m}\right)^2\right] \tag{2}$$

where m is the mean and s is the standard deviation of t.

The mean, median, and dispersion parameters are related as follows:

$$\ln(m) - \ln(t_m) = (1/2)d^2 \tag{3}$$

The cumulative probability distribution function, $F(t)$, is the integral of the probability density function defined by the following:

$$F(t) = \int_0^t f(x)dx \tag{4}$$

TABLE 1 STATISTICAL PARAMETERS OF DIFFERENT VIEWING TASKS

Viewing Task	Mean Value m (Sec)	Standard Deviation s (Sec)	Median Value t_m (Sec)	Dispersion Parameter d (Sec)
Radio	1.44	0.50	1.27	0.376
Left-View Mirror	1.10	0.30	1.06	0.268
Rear-View Mirror	0.75	0.36	0.68	0.455
Speedometer	0.62	0.48	0.49	0.685
Temperature Gauge	1.10	0.52	0.99	0.449
Defroster	1.14	0.61	1.01	0.502
Roadway Name	1.63	0.80	1.46	0.465

TABLE 2 PERCENTILE VALUES OF DRIVER GLANCE DURATIONS IN SECONDS

Viewing Task Percentile	Radio	Left-View Mirror	Rear-View Mirror	Speedometer	Temperature Gauge	Defroster	Roadway Name
5th	0.68	0.68	0.32	0.16	0.48	0.44	0.68
10th	0.78	0.75	0.38	0.20	0.56	0.53	0.81
15th	0.86	0.80	0.42	0.24	0.62	0.60	0.90
20th	0.93	0.85	0.46	0.28	0.68	0.66	0.99
30th	1.04	0.92	0.53	0.34	0.79	0.77	1.14
40th	1.16	0.99	0.60	0.41	0.89	0.89	1.30
50th	1.27	1.06	0.68	0.49	0.99	1.01	1.46
60th	1.40	1.13	0.76	0.58	1.11	1.14	1.64
70th	1.55	1.22	0.86	0.70	1.26	1.31	1.87
80th	1.74	1.33	0.99	0.87	1.45	1.53	2.16
85th	1.88	1.40	1.08	1.00	1.58	1.69	2.37
90th	2.06	1.49	1.21	1.18	1.77	1.91	2.66
95th	2.35	1.65	1.43	1.51	2.08	2.30	3.14

The characteristic features of the lognormal probability density function are described by Aitchison and Brown (*6*) and by Ang and Tang (*7*). This function has previously been applied to represent the probability distribution of brake reaction times of drivers (*8,9*). The skewness, or lack of symmetry, of this analytical model closely approximates the skewness exhibited by experimental data of driver glance durations published in the literature.

Because this function is a two-parameter model, one of the three statistical parameter estimates of mean, median, and standard deviation can be evaluated if two are known. The percentile estimates of driver spare glance durations for the seven visual tasks investigated are presented in Table 2.

CLOSENESS OF FIT TO EXPERIMENTAL DATA

The Kolmogorov-Smirnov test was used to test the goodness of fit between the lognormal analytical model and the experimental data of radio glance times reported by Rockwell (*5*). In this test the actual cumulative probability distribution function, $S_n(t)$, of the experimental data is computed for different values of successive driver glance durations. The corresponding theoretical cumulative distribution function, $F(t)$, is then computed for the same glance durations, and the absolute value of the difference between these values is denoted by $D(t)$, where $D(t) = F(t) = - S_n(t)$.

For the radio glance durations shown, the number of observations was 1,250. The critical value of D, denoted by D_{cr}, at the 5 percent level of significance, is given by the following:

$$D_{cr} = \frac{1.36}{(1{,}250)^{1/2}} = 0.0385 \qquad (5)$$

The successive values of $D(t)$ are presented in Table 3. It is clear that the maximum value of $D(t)$ is $D(0.70) = 0.0321$, which is less than the critical value of 0.0385. This test, therefore, confirms the hypothesis that the lognormal probability density model chosen accurately fits the radio glance data published by Rockwell (*5*) at the 5 percent level of significance.

The Kolmogorov-Smirnov test is considered more powerful than the chi-square test also used in analytical model fitting and is recommended by Ostle and Mensing (*10*). Its advantage is that it does not lump data into discrete categories but compares data in an unaltered form. It is an exact test for all sample sizes.

CONCLUSIONS

The statistical distributions of driver spare glance durations of six in-vehicle and one out-of-vehicle viewing tasks have been generated. The lognormal probability density function has been used to model driver glance durations and statistical parameter estimates computed for each task. A goodness-of-fit test was conducted for car radio glance times, and the fit was found to be satisfactory at the 5 percent level of significance.

The key parametric statistic presented in Table 2 is the 85th-percentile driver glance durations, because this statistic corresponds to the current design driver standard. This value ranges from 1.0 sec in glancing at the speedometer to 2.37 sec in reading roadway name signs. The 85th-percentile design driver completed all other glance times within a 2.00-sec time period.

Of the seven types of viewing tasks studied, two are required for navigation purposes, whereas the other five need only be glanced at occasionally, usually under light to moderate traffic conditions. The two tasks required for navigational purposes are viewing the rearview and side mirrors. The 85th-percentile driver requires 1.08 sec to glance at the rearview mirror and 1.40 sec for the left-side mirror. Both these tasks are, therefore, completed by the design driver within 1.50 sec, which appears to be the maximum recommended time interval that drivers should direct their visual attention away from the roadway scene directly ahead of the moving vehicle. Rockwell (*5*) noted that most drivers attempt

TABLE 3 KOLMOGOROV-SMIRNOV TEST APPLIED TO RADIO GLANCE DATA (*5*): CUMULATIVE PROBABILITY DISTRIBUTION VALUES

Glance Time t (SEC)	Model Estimate F(t)	Data Value $S_n(t)$	Difference $D(t) = \lvert F(t) - S_n(t) \rvert$
0.5	0.0184	0.0028	0.0156
0.6	0.0230	0.0064	0.0166
0.7	0.0565	0.0244	0.0321
0.8	0.1095	0.8080	0.0287
0.9	0.1799	0.1608	0.0191
1.0	0.2625	0.2584	0.0041
1.1	0.3512	0.3616	0.0104
1.2	0.4401	0.4576	0.0175
1.3	0.5247	0.5316	0.0069
1.4	0.6023	0.6008	0.0015
1.5	0.6709	0.6480	0.0229
1.6	0.7305	0.7264	0.0041
1.7	0.7810	0.7688	0.0122
1.8	0.8232	0.8108	0.0124
1.9	0.8580	0.8428	0.0152
2.0	0.8865	0.8708	0.0157
2.1	0.9094	0.8968	0.0126
2.2	0.9281	0.9220	0.0061
2.3	0.9429	0.9396	0.0033
2.4	0.9547	0.9536	0.0011
2.5	0.9642	0.9652	0.0010
2.6	0.9717	0.9740	0.0023
2.7	0.9776	0.9830	0.0054
2.8	0.9823	0.9920	0.0097
2.9	0.9859	0.9976	0.0117
3.0	0.9889	1.000	0.0111

to keep their spare off-roadway glance durations below 1.5 sec despite poor in-vehicle display designs and poor roadway designs. These drivers apparently prefer to take several short glances to a target rather than one or two longer glances, which enables them to drive more safely. Drivers also appear to adjust to denser traffic conditions by shortening their spare glance duration times as intervehicular spacings become shorter. Driver glance times appear to decrease with headway spacing. Three in-vehicle display viewing tasks required correspondingly longer glance times by the 85th-percentile design driver. The viewing times were 1.58 sec for the temperature gauge, 1.69 sec for the defroster, and 1.88 sec for the radio. All of these tasks were therefore completed in 1.90 sec by the design driver. These tasks required visual discrimination as well as visual detection, resulting in correspondingly longer glance durations. Because these tasks are usually performed under light traffic conditions, a 1.9-sec value for these glance durations may be acceptable for most drivers.

The design driver needs 2.37 sec to view roadway name signs outside of the moving automobile. This task requires greater visual discrimination than any of the other viewing tasks studied. The glance duration depends on the number and lengths of the names on the roadway sign. The roadway sign is also the only target that is moving in the driver's visual field. Although experimental data on this viewing task are sparse, it appears that a 2.4-sec viewing time for this task is probably too long and may indicate an unsafe driving situation, because 2.00 sec appears to be an upper bound for the time interval that a driver can safely divert vision away from the roadway scene ahead.

The applicability of using the lognormal probability density function to model driver glance durations has been demonstrated. Although current design standards are based on the 85th-percentile driver, it is possible that this standard may be raised at some future date to the 90th- or 95th-percentile driver. Should this change occur, this analytical model can be used to compute these higher driver percentile estimates of driver spare glance durations.

The time spent in viewing the right-side, or far-side, mirror was not studied, because statistical parameter estimates were not available for this viewing task. A rough estimate given by Rockwell (*5*) is that this task takes about 10 percent longer than glancing at the left-side mirror. This estimate would correspond to a value of approximately 1.54 sec for the 85th-percentile driver. However, standard deviation estimates for this viewing task were not available.

It appears that drivers use the right-side mirror much less than the left-side mirror or the rearview mirror. Most driving manuals, such as one issued by the American Automobile Association (AAA) (*11*), recommend the use of both side mirrors and the rearview mirror, as well as direct vision by drivers turning around, before any lane change is made. The AAA manual first advises drivers to check the appropriate mirror to see if there is adequate time and space to perform a lane change safely. It then recommends that drivers turn and look over their shoulder momentarily to view traffic conditions directly before initiating the lane change maneuver. This procedure will eliminate the possibility of not seeing an approaching vehicle because of the presence of a blind spot in the mirror's viewing field. However, Mourant and Donohoe (*2*) found that novice drivers tended to rely primarily on direct over-the-shoulder viewing rather than mirror viewing when changing lanes on a highway.

Nagata and Kuriyama (*3*) found that glance durations to the far-side mirror were longer than near-side mirror glance times. Robinson et al. (*12*) reported that driver mirror glance times ranged from 0.8 to 1.6 sec, with an average value of 1.1 sec. However, they did not break these viewing times down by the types of mirror being used. They tested eight drivers over a 1.5-mi course. The drivers ranged in age from 20 to 25 years, with an average driving experience of 5 years. The average speed on the course was 30 mph, with each driver circling the course about six times. These glance duration estimates agree well with the results derived here and tend to confirm the hypothesis that the lognormal probability model closely fits the experimental data.

ACKNOWLEDGMENTS

The author wishes to acknowledge the assistance he received from the Department of Civil Engineering and the University of Hawaii at Manoa in the preparation of this paper. The research was conducted under a grant from the University of Hawaii Research Council.

REFERENCES

1. R. R. Mourant and T. H. Rockwell. Strategies of Visual Search by Novice and Experienced Drivers. *Human Factors*, Vol. 14, No. 4, Aug. 1972.

2. R. R. Mourant and R. J. Donohue. Acquisition of Indirect Vision Information by Novice, Mature, and Experienced Drivers. *Journal of Safety Research*, Vol. 9, No. 1, March 1977.
3. M. Nagata and H. Kuriyama. Drivers' Visual Behavior with Door and Fender Mirror Systems. *SAE Transactions*, SAE, Warrendale, Pa., 1985.
4. W. W. Wierwille, J. R. Antin, T. A. Dingus, and M. C. Hulse. Visual Attentional Demand on an In-Car Navigation Display System. In *Vision in Driving, Volume II*, Elsevier, Amsterdam, the Netherlands, 1988.
5. T. H. Rockwell. Spare Visual Capacity in Driving Revisited: New Empirical Results for an Old Idea. In *Vision in Driving, Volume II*, Elsevier, Amsterdam, the Netherlands, 1988.
6. J. Aitchison and J. A. C. Brown. *The Lognormal Distribution*. Cambridge University Press, England, 1969.
7. A. H. S. Ang and W. H. Tang. *Probability Concepts in Planning and Design*. Wiley, New York, 1975.
8. G. T. Taoka. Statistical Evaluation of Brake Reaction Time. *Compendium of Technical Papers, 52nd Annual Meeting of the Institute of Transportation Engineers*, Chicago, Ill., ITE, Washington, D.C., 1982.
9. G. T. Taoka. Brake Reaction Times of Unalerted Drivers. *ITE Journal*, Vol. 59, No. 3, March 1989.
10. B. Ostle and R. W. Mensing. *Statistics in Research*, 3rd ed. Iowa State University Press, Ames, 1975.
11. American Automobile Association. *Sportsmanlike Driving*, 9th ed. McGraw-Hill, New York, 1987.
12. G. H. Robinson, D. J. Erickson, G. L. Thurston, and R. L. Clark. Visual Search by Automobile Drivers. *Human Factors*, Vol. 14, No. 4, Aug. 1972.

The contents of this paper reflect the author's view and are not necessarily those of the sponsoring organizations.

Publication of this paper sponsored by Committee on Vehicle User Characteristics.