TRANSPORTATION RESEARCH RECORD 628

Lighting, Visibility, and Railroad-Highway Grade Crossings

TRANSPORTATION RESEARCH BOARD

COMMISSION ON SOCIOTECHNICAL SYSTEMS
NATIONAL RESEARCH COUNCIL

NATIONAL ACADEMY OF SCIENCES
WASHINGTON, D.C. 1977

Transportation Research Record 628
Price $2.40
Edited for TRB by Frances R. Zwanzig

subject areas
03 rail transport
51 highway safety
53 traffic control and operations

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

Notice
The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competence and with regard for appropriate balance.

This report has been reviewed by a group other than the authors according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

The views expressed in this report are those of the authors and do not necessarily reflect the view of the committee, the Transportation Research Board, the National Academy of Sciences, or the sponsors of the project.

Library of Congress Cataloging in Publication Data
National Research Council. Transportation Research Board.
Lighting, visibility, and railroad-highway grade crossings.

(Transportation Research Record; 628)
Includes bibliographical references.
1. Roads–Visibility–Addresses, essays, lectures. 2. Roads–Lighting–Addresses, essays, lectures. 3. Railroads–Crossings–Addresses, essays, lectures. I. Title. II. Series.
TE7.H5 no. 628 [TE228] 380.5'08s [628.9'5] 77-25134
ISBN 0-309-02653-9

Sponsorship of the Papers in This Transportation Research Record

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

Committee on Visibility
Albert Burg, University of California, Los Angeles, chairman
Robert L. Henderson, National Highway Traffic Safety Administration, secretary
Charles W. Craig, Archibald C. Doty, Jr., Warren H. Edman, Ralph A. Ehrhardt, T. W. Forbes, Vincent P. Gallagher, S. A. Heenan, Robert T. Hennessy, Charles H. Kaehn, Antanas Ketvirtis, Ken F. Kobetsky, Ralph R. Lau, Nathaniel H. Pulling, Richard N. Schwab, Richard E. Stark, Frederick E. Vanosdall, Ned E. Walton, Earl C. Williams, Jr., William L. Williams, Henry L. Woltman, Kenneth Ziedman

Committee on Railroad-Highway Grade Crossings
Otto F. Sonefeld, Atchison, Topeka and Santa Fe Railway Company, chairman
Howard C. Palmer, Safetran Systems Corporation, secretary
William D. Berg, Lucien M. Bolon, Archie C. Burnham, Jr., Louis T. Cerny, Janet A. Coleman, John P. Eicher, Jay Golden, Bruce L. Gordon, William J. Hedley, John Bradford Hopkins, Robert C. Hunter, Robert A. Lavette, Richard A. Mather, Patrick Joseph McCue, Hoy A. Richards, Eugene R. Russell, C. Shoemaker, Max R. Sproles, B. M. Stephens, Marshall Suloway, Robert C. Vanstrum, Richard A. Wiita

K. B. Johns, Transportation Research Board staff

Sponsorship is indicated by a footnote at the end of each report. The organizational units and officers and members are as of December 31, 1976.

Contents

Grade-Crossing Warning-System Technology

John B. Hopkins, Transportation Systems Center, U.S. Department of Transportation, Cambridge, Massachusetts

This paper reviews the objectives, content, and results of a large number of research projects sponsored by the Federal Railroad Administration and related to possible improvement concepts associated with motorist-warning systems at railroad-highway grade crossings. The benefits sought included increased effectiveness, reduced cost, and elimination of institutional constraints. The subjects that were investigated include the application of modularization concepts and alternative components in warning-control logic systems, cost reduction in automatic gate equipment, flashing lights using xenon flashlamp technology, functional requirements and the relevant equipment for lightning protection and standby power, and studies of alternative or novel warning system concepts. The potential for meaningful advances is found to be limited and is severely constrained by the technically challenging nature of the functional and safety requirements.

During the past decade an average of nearly 1300 people/year have been killed in collisions between trains and motor vehicles at railroad-highway grade crossings. Although the trend has been steadily downward, the magnitude of this loss, in what would seem to be preventable accidents, has generated increasing concern. One form that this interest has taken is that of federally sponsored research directed toward the technology of automatic warning devices. Although the causes of grade-crossing accidents—estimated at 12 000/year—are basically due to unsafe driver behavior, there is evidence that the installation of modern train-activated motorist warnings can provide dramatic safety benefits. Flashing lights alone typically reduce accident occurrence by 60 to 70 percent, and automatic gates reduce casualties by at least 90 to 95 percent in most cases (1, 2). Yet, little more than one-quarter of the 220 000 public crossings in the United States are equipped with such devices, largely because of the relatively low hazard associated with many of these crossings and the substantial cost of warning equipment. Also, the many accidents at railroad-highway intersections equipped with flashing lights or other active devices show that their effectiveness, as well as their cost, should be improved. Thus, even for this fundamentally nontechnological problem, research relating to warning-system equipment may significantly improve safety, possibly through more credible or more conspicuous warnings. Reductions in system costs could have a similar impact by permitting wider installation of active warnings or by allowing the use of more elaborate or sophisticated systems within a fixed budget. Innovative technology could also offer a means to circumvent institutional constraints on the implementation of protection. Finally, simply delineating and characterizing the performance and attributes of the available hardware permit more effective selection among the alternatives available.

CONTEXT AND SCOPE OF RECENT RESEARCH

The concepts on which current grade-crossing warning systems are based are not new. The train is detected by a century-old principle, the track circuit that drives a gravity-relay logic system. The flashing lights, gate mechanisms, relays, and such that are currently being installed strongly resemble the devices that have been used for decades. This equipment is highly reliable and virtually fail-safe. Over the years it has been continually improved, and recently solid-state electronic circuits and new materials have been used more and more. Very high levels of performance and reliability are demanded by users, and the attainment of these characteristics in severe operating environments has resulted in technically impressive but relatively expensive equipment. The hardware alone commonly costs over $10 000 for a simple flashing-light installation, and the cost at a multiple-track installation with gates or cantilever mountings or both is several times that. (Installation labor is typically an equivalent expense.)

Research efforts by the industry have always been constrained by basic economic considerations. On the one hand, the resources are limited: Until quite recently, the total market for grade-crossing warning equipment was only about $10 million/year, divided among several suppliers. Further, the results of even very successful research may be slow to appear. The high standards of performance and the very long lifetime required of equipment have made railroads very cautious about accepting new devices, and this inclination is increasingly supported by considerations of legal liability. Thus, under the best of circumstances, several years of successful field testing, which must be preceded by the full development of a new device, are necessary before a substantial market can be hoped for. On the other hand, the technical challenge to improve performance or develop lower cost designs without compromising reliability, safety, maintenance requirements, or resistance to the environment makes efforts in this area very costly, and the field-testing phase, which generally requires substantial redesign and refinement, can more than double the original investment. The possibility of failure is always high, and research efforts going significantly beyond creative product engineering are rarely possible or occur very slowly. The technical advances of recent years—audio-frequency overlay equipment, motion-sensitive detection, improved lights, the application of solid-state technology, the use of alternative gate-arm materials, and such—have been carried out almost entirely within the framework of existing concepts and practices and have typically required lengthy periods for completion.

These constraints have permitted significant improvements in system performance, but may limit the application of advances in industrial and aerospace electronic technology to the improvement of grade-crossing warnings. Thus, the Transportation Systems Center, under the sponsorship and guidance of the Office of Research and Development of the Federal Railroad Administration, has recently assessed the potential technical practicality and the economic viability of a wide range of alternative component and system concepts. The dominant goal of these efforts is improved grade-crossing safety, but in some cases this has been sought through cost reduction that would permit more widespread installation of active warnings or through better understanding of the technical and economic characteristics of the various alternatives.

The discussion presented here is limited to that of completed research on crossing-located, train-activated motorist-warning systems. Within that framework several distinctions are made: Advances within a basically conventional, track-circuit context are treated separately from truly innovative system concepts. And, within the former category, there is a natural distinction among control systems, gate systems, lights, and peripheral hardware. [More extensive descriptions of specific studies can be found in a number of reports (3, 4, 5, 6, 7, 8, 9, 10, 11, 12).] This paper comprises only a general presentation of the nature and scope of this wide range of research projects.

CONVENTIONAL EQUIPMENT

Control Subsystems

The potential for significant cost reduction through technical innovation that is still within the basic framework of track-circuit train detection and conventional motorist warnings has been examined in considerable detail (3, 4). The existing equipment and the practices now prevalent, including a detailed analysis of the equipment and installation costs for several types of crossings, were reviewed in terms of the four subsystems that comprise a train-activated warning system: train detection, control, motorist warnings, and interconnections. Within this breakdown, labor and equipment costs were also separated. Labor costs usually represented approximately one-half of the total expense, with about one-quarter of that being shop (rather than field) labor. The individual cost elements showed little potential for cost reduction in areas other than the control subsystem, and even in that area, savings of only 10 to 20 percent appeared attainable, implying a maximum of only 3 to 7 percent impact on the total cost. (Gate arms and drive mechanisms were explicitly excluded from detailed study in this project.)

This project included a process of generating, characterizing, and evaluating concepts by which equipment suppliers might improve their products. Several possible approaches were identified, and their potential benefits were balanced against their greater cost and the uncertainty associated with their realization. The most straightforward path appeared to be that of designing, constructing, and testing systems based on modular components, with different combinations of a small number of basic modules sufficient for simple assembly of the logic and control portions of most installations. In the simplest form, the modules would be based on combinations of existing gravity relays. Other possibly advantageous concepts would extend the modular approach by using alternative, lower cost types of relays or solid-state circuitry in place of the traditional vital relays, which are reliable and fail-safe but costly and physically large. These efforts were conceptual and analytical in nature—no hardware was designed or tested—and firm conclusions are not possible. However, a tentative finding was that the use of the type of signal relay common in Europe, which is highly reliable but based on self-checking rather than on inherently fail-safe design, should seriously be investigated. Mercury-wetted reed relays, which have not previously been used in this type of application, may also merit more detailed examination. Redundancy techniques, which are common in some high-reliability applications, did not appear to be attractive here. For special cases, solid-state components appear to be useful, but constraints relating to market size, temperature extremes, and surge protection make such devices less competitive than might be anticipated; their customary advantages of high operating speed, small size, and sophisticated functional capabilities are not of great value at grade crossings, for which the system logic is relatively simple and time constants of seconds are fully satisfactory.

Automatic Gates

Train-activated automatic gates generally provide a safety effectiveness (accident reduction) of 90 to 95 percent or better. However, their substantial cost has limited their use; a complete crossing using gates (but not cantilever light mountings) typically costs $35 000 to $50 000, compared to $20 000 to $30 000 for flashing lights alone. Although this large differential is partially due to the fact that gates are usually used at more complex crossings such as multilane and multitrack ones, gate hardware itself is not inexpensive and may be 10 to 20 percent of the total cost of the installation. Maintenance costs are also substantial.

Gate breakage is a major problem. Either accidentally or deliberately, it is not uncommon for motor vehicles to drive through lowered gates, or to snag the tip in attempting to go around them. The result is a breakage rate that averages more than 1 gate arm/year for each gated crossing, and may be far greater for particular crossings or localities. This problem may arise from unnecessarily long activation times or from false alarms in which no train reaches the crossing at all, which result in driver annoyance or frustration and lead to the decision to go through the gate rather than wait for it. Or, if the crossing activates just after a large trailer truck has entered it, particularly one that has just made a mandatory stop, the descending gate arm can catch on the vehicle and be damaged. Other sources of damage are vandalism and strong gusty winds. Some of these problems can be alleviated through direct means such as the use of constant-warning-time train detection, but the basic problem remains. Since arms cost $200 to $300 and require substantial installation labor (often at overtime rates), the economic burden on the railroads can be quite significant.

There has been continuing research in this area. The traditional double wooden arm has been challenged by fiberglass and aluminum alternatives, and recently a manufacturer not previously involved in this market has developed a polycarbonate (Lexan) arm. Gate mountings that shear, permitting the arm to drop free rather than break, have come into widespread use. Nonetheless, further improvements through application of recent advances in materials and structures are possible.

A clear distinction between required and desired arm characteristics is difficult to make: A number of features must be balanced within certain constraints. Some railroad personnel find gate-arm breakage useful in the event of an accident, as this provides evidence, which can be brought into court, that the gate was lowered at the time of an accident. Too rigid an arm could cause a derailment if it were to be knocked onto the tracks. An electrically conductive (metal) arm might be a safety hazard if it came into contact with power lines when in a raised position. The costs of providing resistance to initial breakage must be balanced against those associated with ease of repair, unless both can be combined.

The present gate-drive mechanisms have an impressive record of performance and reliability. However, several factors, including limited research resources and the existing industry standards, have limited the range of alternative approaches and components used. The application of recently developed materials, new structural concepts, and improved components might provide significant economic benefits without compromising performance, safety, reliability, or service life.

Accordingly, this area was studied (5, 6) through a sequence of tasks that included a thorough review of existing practices, specifications, and regulations; recommendations concerning areas in which modifications of existing requirements might permit significant overall cost reductions without compromising safety or performance; and the generation of new concepts for gate systems. These concepts were then subjected to as thorough an engineering and economic analysis as possible, and recommendations concerning possible future research and development were made.

Three areas, the gate-drive mechanism, the arm support, and the gate arm, were identified as targets for possible advances. The suggested concept for a low-cost drive mechanism that could offer significant cost benefits if all of its components should prove practical is based on combinations of several commercially available elements:

1. A high-speed sealed motor integrated with a sealed high-ratio, high-output-torque gearbox and output shaft;
2. A sealed ball-bearing unit that uses a long-life lubricant for support of the output shaft;
3. Sealed switches and relays of high reliability; and
4. A compact, weatherproof, lightweight enclosure.

The use of a small motor with a high gear ratio would require special attention to achieve a fail-safe operation (the lowering of the gate in the event of a power outage); this area would require specific and substantial investigation if this concept were to be pursued further.

Another possible approach is the use of a pneumatic drive mechanism, which has a number of attractive features.

Another concept suggested is that of a swing-away gate-arm support that uses a semiflexible arm on a pivoting mechanism. This allows the arm to swing up out of the way when struck and afterward return by gravity to its original position. For the arm itself, the use of new fabrication materials in basically conventional arm structures could offer substantial potential benefits: The material recommended is a phenolic resin-impregnated honeycomb encased in a fiberglass-reinforced polyester tube. An arm of this material used with an effective swing-away resetting mounting would cost substantially less over its potential service life.

Flashing Lights

Conventional grade-crossing flashing-light systems generally reduce accidents by 60 to 75 percent; these devices are by far the most common train-activated motorist warnings now in use. Yet, over one-third of the present fatalities occur at the approximately 50 000 higher traffic-density crossings marked by active warnings. The primary causes of these collisions appear to be motorist inattention, carelessness, misjudgment, error, or inebriation. However, since active devices have a strong positive effect, even these factors could be overcome (7, 13).

There are two usual assumptions concerning the needs of a motorist approaching a grade crossing. The first is that he or she is less likely to be aware of the presence and hazard of the crossing than of that of a normal highway intersection; hence, a high degree of alerting effectiveness is necessary. The second is that, since the situation and the hazard at a crossing are significantly different from those at a crossroad, the warning should be immediately and unequivocally identifiable as being associated with a railroad crossing.

A possibly effective and practical innovation is the use of the high alerting effectiveness and power-conversion efficiency associated with the very short flash duration (less than a millisecond) of xenon (strobe) lamps. It is technically a relatively simple matter to mount conventional xenon lamps in standard grade-crossing flasher heads in place of the normal incandescent bulbs, and to add an appropriate power supply. In the quiescent state the crossing appears essentially the same as with conventional lamps, but, when activated, the red strobe lights supplement the existing incandescent units in a highly alerting fashion. Subjective judgments of the potential safety effectiveness of xenon lights have been highly favorable, and some of their greater power efficiency could be used to provide a wider beam width, which would make them less vulnerable to misalignment.

Peripheral Subsystems

Two peripheral areas, which do not have major direct impacts on system cost but which are serious constraints on system design and reliability, are those of the protection of the equipment from the effects of electrical surges (primarily those associated with lightning) and the provision for emergency or backup power supplies.

Lightning annually damages or destroys millions of dollars worth of railroad signaling equipment in the United States. This problem has existed since the inception of electrical signaling and communication systems on the railroads. Protective devices and techniques have been developed to control this effect, but they do not eliminate it. The situation has become even more serious in recent years due to the increasing costs of repairs (particularly labor) and the introduction of solid-state electronic components, either to replace older electromechanical relay systems or to increase functional capability. Although the solid-state components are often useful, or even necessary for some devices, they are inherently more vulnerable to damage from lightning and other electrical surges than are traditional electromechanical devices. Although there are now available a wide range of surge-protection devices, as well as analytical techniques for better understanding of particular applications, these advances appear to have been used to a lesser extent in railroad signaling than in other areas. Standardized specifications for surge testing, surge resistance, and surge protective devices would also be desirable, and there would be a marked improvement after greater involvement of trained surge-protection specialists in the problems of railroad signal equipment.

A survey (12) of the requirements and technology relating to the provision of standby power to operate crossings in the event of failure of the commercial or other 110-V line power found that the railroads voluntarily assume far more rigorous standards in this matter than are imposed by public bodies, often requiring sufficient battery capacity to last between scheduled maintenance visits (1 to 2 weeks). Batteries with regulated charging units appear to be the preferred approach although there are a number of alternatives that may have advantages in certain situations. Solar power, for example, is beginning to find a small but significant role in powering railroad signal systems.

INNOVATIVE SYSTEM CONCEPTS

The relatively high interest in recent years in grade-crossing safety has generated many suggestions concerning possible improvements. Often, these ideas reflect a misunderstanding of the technology involved, the functional requirements of the system, acceptable economics, or accident-causal factors, but other ideas justify more thorough consideration. A study of the technical feasibility and potential benefits of truly innovative system concepts for train-actuated motorist-warning systems (14) is discussed below.

Communication Link

The track-circuit systems used almost universally in the United States detect train occupancy at any point in the signal block. A possible alternative is the detection of trains only at entrance and exit points, with that information then communicated to a central storage and processing point. This is a common practice in Europe. Two communication-link methods are possible: microwave telemetry, for which good power efficiency can be obtained by the use of tightly focused beams (8), and very-high-frequency radio transmission, which is not limited to line-of-sight operation (8, 9). It may be that each method has a role to play, depending on particular circumstances; in any event, this affects only a small part of the system design and performance.

The principal technical difficulty with this overall concept is the selection of a train-detection device that meets all of the requirements of reliability, fail-safe operation, low power consumption, long lifetime, invulnerability to extreme environments, and low cost. At least one existing sensor appears to be adequate but costs over $1000 (several would be required) and must be attached to the tracks. A variety of physical principles that might lead to a lower cost device have been identified, but any serious development effort would inevitably be lengthy, expensive, and of uncertain outcome, since this element is at the heart of safe system operation and must meet very rigorous performance and reliability standards.

However, the potential cost reduction of the communication-link approach appears to be limited, and there is a definite possibility that its price would ultimately exceed that of conventional systems. Its functional advantages are also uncertain: The use of speed-sensitive train-detection devices might facilitate constant-warning-time operation, but there is considerable system complication when one attempts to equal the performance of conventional motion-sensing equipment. The major advantage of this concept is the possibility of realizing a system that could be operated by a public authority, such as a highway department (15). The use of track circuits now involves the railroads so intimately that this is virtually impossible. However, the full benefits of this course of action would require retrofitting many of the more than 50 000 crossings that now have conventional train-activated warnings.

Radar Train Detection

The use of crossing-located radar has been a popular idea for several years, partially stimulated by the availability of simple solid-state radar modules and apparently analogous uses in motor-vehicle speed monitoring, small-boat safety, and military-perimeter surveillance. However, as a train-detection method, this concept has numerous weaknesses, particularly in line-of-sight restrictions, the absence of fail-safe operation, and inadequate performance at multiple-track crossings or those near parallel highways. In addition, even the partial satisfaction of the necessary rigorous specifications escalates costs to an unacceptable level. While such an approach might someday be practical as a parallel subsystem for providing constant-warning time, it is unlikely to be viable as a primary means of train detection.

Track Radar

One new form of track circuit that appeared promising is significantly different from present track circuits in that it does not rely exclusively on circuit characteristics, but on the return reflection of audio-frequency electrical signals transmitted down the track from the crossing. Through the use of correlation circuits the elapsed time from the origination of the signal to the receipt of its reflection can be measured to give a precise indication of the location of the reflecting element—typically a train, short-circuiting the rails. The velocity and direction of the movement of the train can in principle be determined by following its location as a function of time, so that the constant-warning time can theoretically be obtained. This system should permit automatic compensation for changes in the electrical properties of the track and ballast. However, the basic principle of operation has not been demonstrated in practice, or even analyzed in depth, and the electrical variability of typical track structures will undoubtedly pose difficulties. The sophisticated equipment required might make the method too expensive to achieve.

Locomotive-Mounted Transmitters

Another frequently suggested concept involves devising a means by which a locomotive can signal the crossing of its impending arrival. A number of variations are possible and lead to systems that differ widely in cost and probable performance. The simplest variation might be a continuously operating locomotive transmitter that activates warnings at all crossings within range. The next level of complexity would provide for transmission only when the train is approaching or occupying a crossing, with activation of the transmitter occurring through manual means or some wayside device in advance of the crossing. Far more elaborate concepts could also be generated; for example, a locomotive might have an odometer to monitor its exact location continuously, with the transmissions coded for particular crossings and activated from a route specification stored in a microprocessor memory, which could also accommodate the train speed.

There are weaknesses inherent in all of these concepts. In general, they require the sacrifice of the fail-safe principle that has guided railroad signal practices for many decades. In them, the normal condition at the crossing—no signal received—is no different for the case of the approach of an unequipped or malfunctioning locomotive than it is for no train approaching. Although this difficulty could be circumvented by operating rules and engineer intervention, it is still a non-safe-failure mode. In addition, there is the serious limitation that all locomotives and other power units would have to be appropriately equipped for the system to be totally effective and safe at even a single crossing. Thus, widespread application at crossings would be required to justify the modification of the locomotive fleet.

A simple device that activates crossing signals indiscriminately might generate so many unnecessary or excessively lengthy advance warnings as to lose all credibility with motorists, and even this type of system could be relatively costly when fabricated to high standards of equipment reliability. On the other hand, dependence on the train crew to deal with all cases, or merely the exceptions, imposes an additional burden on them and introduces a major potential point of controversy in the event of an accident.

One can envision a system in which a locomotive arriving at a crossing interrogates a passive wayside device, receives a coding for that crossing, and transmits the appropriate signal to activate that crossing only. This system could also incorporate on-board speed sensing and alter the warning as required. A specific sensor would be needed at the crossing to deactivate the warning after the entire train had passed. It would also be possible to require that a crossing-located transponder answer the locomotive when the signals were activated; the failure to respond in a brief interval would alert the engineer to the possibility that the crossing warnings had not been activated. However, these more elaborate concepts will be expensive, which will limit their attractiveness as compared to conventional warning systems. On the other hand, the simpler, lower cost systems will not be as effective in warning the motorist at the crossing. The most reasonable concept is probably an essentially manual approach in which the engineer activates the transmitter when necessary and attempts to stop if no confirmation is received, and this is combined with a very simple, low-cost warning such as a single strobe light at the crossing. It would be important that motorists be able to distinguish between this possibly unreliable indication and the standard fail-safe signals, to avoid any diminution of the effectiveness of the latter. An approach of this type might be only slightly inferior in safety effectiveness to conventional equipment, but at a far lower cost. The major use of such a warning system would be for crossings with lower traffic volumes that do not justify the expense of current systems. However, it is not clear that this type of compromise is attainable and cost-effective, nor that it would be generally acceptable to safety authorities or the public.

Train Indicator

Under various specialized circumstances, both in Europe and in the United States, train indicators, wayside signals to the locomotive engineer that forbid entry into a crossing until the motorist-warning system has been activated, have been used. This broad concept can be implemented in many ways, but often involves a fail-safe arrangement in which the wayside signal (train indicator) is normally red and changes to green only when the crossing signal is on. The typical use is on heavily traveled roads that are crossed by infrequent, short, low-speed trains that can easily stop if necessary; in some cases the normal procedure is for the train to stop so that the train crew can manually activate the warnings. However, this concept could be used in a way that might permit significant reductions in the cost of warning equipment. At sufficiently low speeds, even a moderately large freight train can stop within the normal grade-crossing-approach circuit distance if the crossing warning does not activate. Even for a relatively long, heavy train, this condition would be satisfied at train speeds of up to approximately 32 km/h (20 mph). For shorter trains, or a longer than normal prearrival warning time at the crossing, the situation is still less restrictive. Thus, this approach could be used for crossings for which factors such as track conditions limit speeds sufficiently. This is not to imply that trains should be stopped at crossings or that speed limits should be reduced. In addition to the havoc this could play with schedules and operating costs (there is approximately 0.6 public crossing/railroad route-kilometer in the United States, with 1 in 5 marked by train-warnings), such an approach would lead to trains moving over crossings slowly or from a dead stop, which would generally increase exposure time (and thus hazard) and highway congestion. However, by the use of such systems, the very high equipment-reliability standards now imposed might be relaxed. Since the last increment of reliability is typically very expensive to achieve, this might be a way to reduce cost with no appreciable loss of safety and a very low probability that a train would, in fact, have to stop.

This concept is highly speculative. Its viability depends on the number of crossings to which it might apply and their accident potential. It further makes the unproven assumption that major cost reductions could be achieved with no more than a very small diminution of system reliability and no increase in maintenance requirements. Nonetheless, it could be of interest for the large number of crossings that are not hazardous enough to justify the expense of conventional warning systems but are the locations of a significant number of accidents.

CONCLUSIONS

The goal of achieving significant cost reductions and increased safety effectiveness at railroad-highway grade crossings purely through technical innovation appears to be attainable only to a limited degree and is technically very challenging. No concepts that offer dramatic improvement in system economics have been generated, although a number of avenues that might lead to significant, if modest, cost savings have been identified. These are primarily improvements in gate arms and drive mechanisms and in control-system modularization, possibly using logic elements other than gravity relays.

A promising potential advance in safety effectiveness is in the use of xenon flashlamps (strobe lights) at grade crossings, where the conspicuity and alerting impact of the short-duration flashes would be a significant improvement. This would also generate economic benefits through the elimination of the requirement for extremely stable (and expensive) cantilever mounting structures that are more costly than gates.

A high degree of consistency in the advance warning time of crossing signals enhances their credibility with motorists and therefore improves safety. However, none of the suggested alternative methods clearly represents an improvement over current techniques, and this area is sufficiently important to justify further investigation.

The existing technical barriers to the transfer of the operational responsibility for crossing warnings could be overcome with a sufficiently strong motivation. However, the required development effort would be substantial, and the equipment cost is unlikely to be lower. Thus, the ultimate attractiveness of this course depends on factors and judgments that are not appropriate to this discussion.

The available array of warning systems lacks equipment that is sufficiently low in cost for truly widespread installation, even at crossings of quite low hazard, but the development of this equipment is unlikely to be achieved without some diminution of safety effectiveness, a consequence that has generally not been considered acceptable. Concepts such as those of a simple locomotive-mounted system or a train indicator could offer substantial overall safety benefits, but raise serious questions of policy in public-safety matters.

These conclusions are only tentative, partially subjective, and based on the present understanding of the causes of the problem and of relevant technologies. Should the potential advantages be sufficiently attractive, further research would be necessary to confirm and define more precisely the magnitudes and values of possible improvements. Such decisions must be made within the context of the practicality and acceptability of new approaches to a system that is now structured around particular safety requirements, technology, skills, inventory, and maintenance standards. The development of the actual equipment would be a lengthy, expensive, and inevitably somewhat speculative endeavor. On the other hand, the range of possibilities and concepts that has been identified by the research described here suggests opportunities and may stimulate significant advances in this long-standing problem of public safety.

ACKNOWLEDGMENT

No claim of originality is made for many of the concepts here identified, and some may even be protected by existing corporate patents. The judgments expressed here represent in part the personal opinions of the author, deriving from experience and professional judgment, but nevertheless are not to be taken as final or absolute. In particular, these comments do not necessarily represent the views or policies of the U.S. Department of Transportation, the Federal Railroad Administration, or the Transportation Systems Center.

REFERENCES

1. D. Schoppert and D. Hoyt. Factors Influencing Safety at Highway-Rail Grade Crossings. NCHRP, Rept. 50, 1968.
2. The Effectiveness of Automatic Protection in Reducing Accident Frequency and Severity at Public Grade Crossings in California. California Public Utilities Commission, June 1974.
3. C. L. DuVivier, L. M. Rogers, W. Sheffeld, and W. J. Foster. Potential Means of Cost Reduction in Grade-Crossing Motorist-Warning Control Equipment. Federal Railroad Administration, Rept. FRA-ORD-75-80, Vol. 1, 1977.
4. F. R. Holmstrom. Potential Means of Cost Reduction in Grade-Crossing Motorist-Warning Control Equipment. Federal Railroad Administration, Rept. FRA-ORD-77-37, Vol. 2, 1977.
5. A. St. Amant. Potential Means of Cost Reduction in Grade-Crossing Automatic Gate Systems. Federal Railroad Administration, Rept. FRA-ORD-77-06, Vol. 1, Feb. 1977.
6. J. Duttera and M. Friedland. Potential Means of Cost Reduction in Grade-Crossing Automatic Gate Systems. Federal Railroad Administration, Rept. FRA-ORD-77-06, Vol. 2, Feb. 1977.
7. J. Hopkins and E. White. Improvement of the Effectiveness of Motorist Warnings at Railroad-Highway Grade Crossings. Federal Railroad Administration, Rept. FRA-ORD-77-07, Feb. 1977.
8. J. Hopkins and others. A Communication-Link Approach to Actuation of Grade-Crossing Motorist-Warning Systems. Federal Railroad Administration, Rept. FRA-ORD-75-80, July 1975.
9. F. H. Raab, M. C. Brooker, T. E. Ryan, and J. R. Waechter. Innovative Concepts and Technology for Railroad-Highway Grade-Crossing Motorist-Warning Systems. Federal Railroad Administration, Rept. FRA-ORD-77-37, Vol. 1, Aug. 1977.
10. D. D. Peterson and D. S. Boyer. Innovative Concepts and Technology for Railroad-Highway Grade-Crossing Motorist-Warning Systems. Federal Railroad Administration, Rept. FRA-ORD-77-37, Vol. 2, Aug. 1977.
11. F. R. Holmstrom. Lightning and Its Effects on Railroad Signal Circuits. Federal Railroad Administration, Rept. FRA-ORD-76-129, Dec. 1975.
12. F. R. Holmstrom. Standby Power for Railroad-Highway Grade-Crossing Warning Systems. Federal Railroad Administration, Rept. FRA-ORD-76-286, Sept. 1976.
13. J. Hopkins and F. R. Holmstrom. Toward More Effective Grade-Crossing Flashing Lights. TRB, Transportation Research Record 562, 1976, pp. 1-14.
14. J. Hopkins and M. Hazel. Technological Innovation in Grade-Crossing Protective Systems. Transportation Systems Center and Federal Railroad Administration, Rept. DOT-TSC-FRA-71-3, June 1971.
15. J. Hopkins. Technological Aspects of Public Responsibility for Grade-Crossing Protection. TRB, Transportation Research Record 514, 1974, pp. 33-43.

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

Traffic-Control Measures at Highway-Railway Grade Crossings With Provisions for Light Rail Transit

J. Schnablegger, Edmonton Transportation Planning Branch
S. Teply, Department of Civil Engineering, University of Alberta

Railway rights-of-way in cities are attractive alternatives for transit corridors, but, for modes that are not fully grade-separated, such as light rail transit systems, there may be problems with combined railway and transit crossings of arterial streets. This situation has been studied in Edmonton, Alberta, where a light rail transit line is under construction. The surface portion of this line is along the railway right-of-way, and as a result, the operation of its eight grade crossings is regulated by railway authorities. The short headways of light rail transit could cause frequent disturbances to the road traffic that operates at saturation during peak hours. This paper illustrates the method used for the analysis of the problem and discusses the surveys conducted. The basic principles governing the solutions to the grade-crossing problem are (a) the coordination of adjacent signalized intersections in such a way that the impact of the crossing closure is minimized and the system recovers shortly after the closure, (b) the integration of light rail transit scheduling and control with traffic control, i.e., restricting the closures to the periods of minimum impact on road traffic, and (c) the use of special features to increase safety.

The northeast sector of Edmonton contains industrial and recreational complexes and has a residential population of approximately 100 000 persons, which is expected to increase to 150 000 persons by the year 1980. One-third of this growth is expected to occur in new outlying areas, and the balance will be in the presently developing areas and the older developed areas.

At present, the transportation needs of the area are served by an arterial road network and the public-transit system of buses and trolleybuses. To serve the future needs of the area, the construction of a light rail transit (LRT) line augmented by a feeder bus system was approved in 1973. The LRT system is suitable for the population thresholds expected, and the availability of the Canadian National Railway (CNR) right-of-way along which the line can run makes it a cost-effective option.

The line, which will be in operation by 1978, will be 7.2 km (4.5 miles) long, 1.6 km (1 mile) of which will be tunneled beneath the streets of downtown Edmonton. The remainder will operate on the surface along the CNR right-of-way. Two stations will be underground and three will be on the surface (Figure 1). So that the line will be cost-effective, the surface portion will, at implementation, retain eight existing grade crossings.

At present, the arterial roadways in the area operate at a high level of service during peak hours and special events, and extensive queues on the links crossing the railway tracks are common. The introduction of the LRT line will increase the disruptions of these arterial roadways, and this loss of capacity and the decreased safety will be potential disruptions to the LRT operations. Thus, to achieve safe and efficient transportation in the northeast sector of Edmonton will require integrated management of all modes including the LRT.

PROBLEMS

Existing Situation

At present 20 to 24 railway trains traverse the grade crossings in a 24-h period, but since the majority of them do so during off-peak periods, they are not a major traffic disruption.

Nevertheless, the signalized intersections adjacent to the railway crossing [those 35 to 122 m (115 to 400 ft) from them] are a source of serious capacity problems in the morning and afternoon peak hours for the following reasons:

1. Conflict between the major traffic flows from generators north and east of the central business district,
2. Heavy left-turn movements that require 2½ or 3-phase control,
3. Isolated vehicle-actuated operation of traffic signals,
4. Physical restrictions that prevent intersection improvements, and
5. Restrictions to the road network because of the presence of the railway, major industrial and recreation facilities, and topography.

These capacity problems and the directional nature of the traffic cause long queues.

The queueing and capacity problems were surveyed and analyzed by the use of helicopter and surface crews. The major objective of the surveys was to obtain data with which to illustrate the operation of the transportation network in this area. These data were then used as the basis for an analysis of the situation that is expected after the introduction of the LRT. The surveys also showed the interaction of the traffic-actuated signals at two adjacent intersections (Figures 2, 3, and 4).

The schematic example of a real-time-space diagram in Figure 2 shows the degradation of the network performance in area C in the afternoon. The critical traffic conditions develop between 4:15 and 5:30 p.m. as the traffic inflow exceeds the capacity of the downstream intersection. [For clarity, the traffic conditions in the opposite direction are illustrated on a separate diagram (Figure 3).] The inflow traffic is generated at the upstream three-phase intersection during two signal phases with some right turns on red and is discharged at the downstream three-phase intersection during one signal phase. The solid horizontal lines in the diagram show the length of the queue for each traffic lane during three time profiles of the vehicular (green) interval: the beginning, the midpoint, and the end. The actual cycle time for both intersections is identified. The following observations can be made.

1. During the off-peak period, each intersection had a different cycle length. The upstream intersection consistently used a shorter cycle length than did the downstream intersection. The individual offsets varied.
2. As the traffic volumes increased, the cycle length at the upstream intersection increased. Since traffic at the downstream intersection approached saturation, the discharge phase operated at maximum capacity. The offset began to stabilize at this point.
3. During the peak period, two phases of each intersection (west and south approaches) became saturated and operated at maximum capacity. Uniformity of cycle lengths was established and small variations in cycle lengths that were shorter than the maximum were caused by the third unsaturated phase.

Several surveys taken on different days confirmed the consistency of the traffic events illustrated in these diagrams.

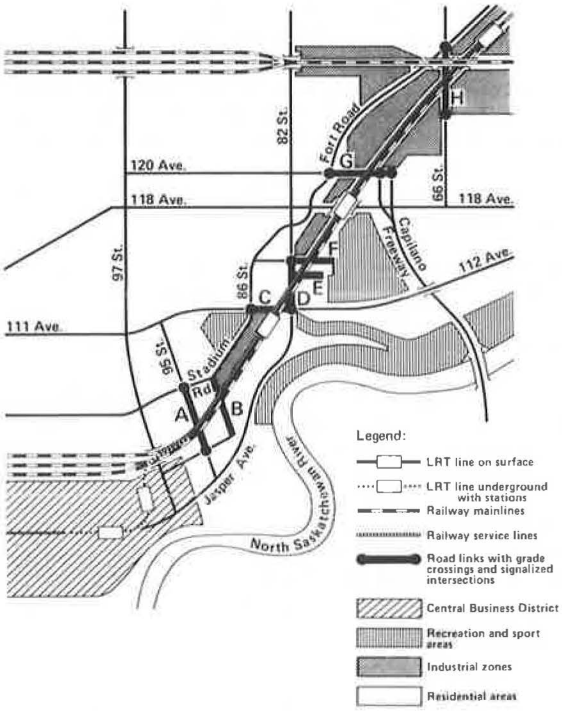
Traffic conditions during the morning peak periods are less severe (Figure 4), but the critical problem of queuing across the track area is still present. The analysis of the diagrams and the helicopter film indicate the following:

1. The operation of two traffic-actuated signals 180 to 360 m (525 to 1200 ft) apart under directionally pronounced saturated traffic flows became similar to a fixed-time-linked system of operation.
2. The traffic flow was the medium that induced the linkage. Because of the deficient capacity of the downstream intersection, this linkage did not produce progression.
3. The approaches to intersections having traffic-flow rates lower than saturation increased the delays and the number of stops in the major directions. These approaches also lowered the overall intersection capacity by excessive extensions of their vehicular (green) intervals, which operated at low levels of service.
4. Because of the width of the railway crossing (two to six tracks) and the queuing phenomena, the number of vehicles that stopped in the crossing area was high (on the average there was one stop longer than 20 s in every second cycle).
5. The average saturation-flow rate at the downstream intersection was 1530 passenger automobile units/h of green time per lane.

Problems Expected as a Result of Light Rail Transit Operation

The existing traffic conditions during peak hours are far from satisfactory, and the introduction of the LRT line will further increase the problems. The LRT trains will operate on a 5-min headway (300 s) in each direction during peak hours and, on the average, will interrupt traffic every 2.5 min. The occurrence of these interruptions will depend on the location of the crossing, the detailed LRT schedule (Figure 5), and the schedule adherence of the trains.

Figure 1. Northeast sector of Edmonton with light rail transit line.



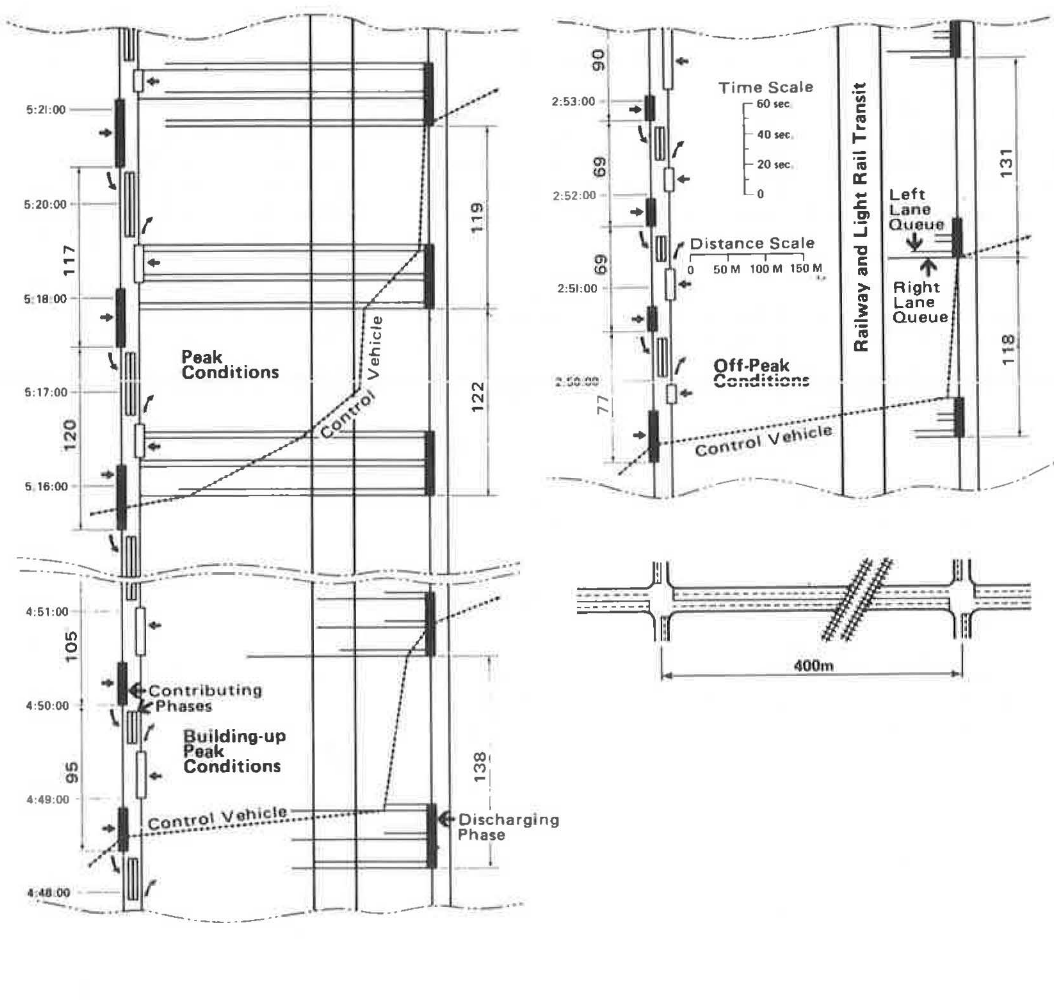
The capacity losses due to these interruptions could be considerable. The cause of these losses is illustrated in Figure 6. This time-space diagram shows the hypothetical trajectories of vehicles and the traffic shadow of a road closure at the railway crossing in a fixed-time system. Such shadows have a specific probability of occurrence. If a normal distribution of offsets between the railway crossings and intersection 2 downstream (both operating in the traffic-actuated mode) is assumed, the capacity losses are estimated to be 10 percent for simple configurations and lighter traffic conditions (e.g., area A) and up to 25 percent for more complex configurations and heavier traffic (e.g., area C).

The increased delays and the increased number of stops are also illustrated in Figure 6. The stopping occurs in front of the railway crossing because of the closure and because of the oversaturation of the downstream intersection. During the peak periods the delays and the number of stops are a function of the capacity losses. In off-peak periods, there will still be significant delays and stops because the tangential routes crossing the LRT line still carry traffic volumes at close to saturation.

The safety hazard will increase as both queuing and train frequency increase.

There will be irregularities in the LRT schedule if the safety problem is resolved by applying restrictions to the LRT operations, such as reducing speeds or inducing stops to allow queues to clear the tracks.

Figure 2. Schematized sections of real-time-space diagram for afternoon westbound traffic at railway crossing C.



Design Objectives

The design objectives to solve the problems were specified as follows:

1. The basic objective is the safe operation of both the LRT and the automobile traffic.
2. The related objectives are the minimization of the following: (a) disturbances to the LRT schedule, i.e., delays caused by disruptions at the crossing; (b) capacity losses to the road network; (c) delays to the road traffic; (d) the number of forced stops to the LRT trains; and (e) the number of stops to the road traffic.

TIMING REQUIREMENTS FOR THE CONTROL OF RAILWAY GRADE CROSSINGS

One of the elements that affects the objectives outlined above is the length of the closures of the railway crossings. In Canada there are no regulations that are specific to the operation of LRT vehicles. However, because the Edmonton LRT line used the CNR right-of-way, railway jurisdiction applies, and unfortunately, these regulations do not recognize the performance features of LRT technology.

In Canada the operation of the railways is under the jurisdiction of the Rail Transport Committee of the Canadian Transport Commission. The regulations for grade crossings are in General Order Number E-6 of the Board of Transport Commissioners for Canada.

Section I, paragraph 8 (1) of this order requires that crossing signals operate for not less than 20 s before the crossing is entered by a train at a speed in excess of 16 km/h (10 mph) and that, if the roadway distance between the governing signal and the clearance on the opposite side of the farthest protected track is more than 10.7 m (35 ft), the operating time of 20 s be increased 1 s for each additional 3 m (10 ft). Signals must continue to operate until the train has cleared the crossing. Paragraph 12 identifies gates as adjuncts to signals. The requirement for gates is a function of train and vehicular traffic.

To illustrate the differences between the timings required by the treatment of the crossing according to railway regulations and according to the rules for a signalized traffic intersection, three timings (railway crossings without gates, railway crossings with gates, and LRT crossing as a signalized traffic intersection) are shown in Figure 7.

The most efficient operation would clearly be case 3, the signalized intersection. The governing regulations, however, require the use of gates and flashing lights. The use of separate controls, one for the LRT and the other for the CNR trains, was rejected because of the hazards of dual indications for vehicular traffic. Thus, it was decided to design a control system that uses railway gates and flashing lights with some timing allowances granted by the Canadian Transport Commission for the LRT operation. In addition, the control logic was designed so that it could use the operational features of the LRT.

Figure 8 shows the locations of the LRT crossing-control and detection equipment. The following table gives examples of the sequences of events and the associated timings for one LRT train and for the extreme case of two trains traveling in different directions.

Action	Cumulative Elapsed Time (s) 1 Train	2 Trains
Train detection	0	0
Signals start flashing	1	1
Stop signal for LRT changes to proceed	2	2
Trip stop deactivates	2	2
Gates start closing	6	6
Gates fully closed	19	19
Train enters crossing	21	21
Train in opposite direction enters extended detection	–	30
Train clears crossing	31	31
Gates start lifting	31	–
Gates fully upright	38	–
Detection of train in opposite direction	–	47
Stop signal changes to proceed	–	49
Trip stop deactivates	–	49
Second train enters crossing	–	68
Second train clears crossing	–	78
Gates start lifting	–	78
Gates fully upright	–	85

The detection of the LRT vehicles is achieved through track circuits. If the controls fail to respond, the LRT-system signals maintain a stop indication, and if the train violates this signal, emergency braking is applied to stop the train before it reaches the crossing.

Under normal railway practice, if, shortly after a train has left a crossing, another train is detected coming from the opposite direction, the gates lift and lower again in a short sequence. Because of the frequency of LRT movements this is not desirable. The extended detection circuit prevents this and maintains a minimum time of 10 s between sequential gate closures to allow for road traffic.

CNR crossing control can be incorporated into the system to achieve consistent protection of the crossing. Because the rolling stock used by the railway is unable to operate in the same manner as that of the LRT, only the detection and extension features can be incorporated, but, by using these features, a railway train can extend an approach circuit for LRT and vice versa. This will result in a safer and more consistent operation than if the railway control were not integrated.

DESIGN PRINCIPLES

The basic traffic-control philosophy for the areas and intersections being discussed is the development of a system that could recover after the disruption caused by the LRT crossing closure. To implement this philosophy, three major principles for the design of controls were adopted.

The first principle is the coordination of the traffic signals so that extensive queuing across the railway crossing can be eliminated. This can be done by controlling the capacity of the upstream signals that feed this link so that the queuing in front of the downstream intersection is reduced to an acceptable length, and vehicles that arrive subsequently will then move through the downstream intersection without stopping (Figure 9).

At the same time, this measure will reduce the number of stops and delays in the system. In most cases, vehicles will be stopped only on the approaches to the upstream intersections and will move through the system on a green wave.

The second principle is the integration of the operation of the traffic signals with the LRT controls. The objective of this is to use the periods of time provided by the shadow of the red signals at adjacent intersections for the LRT crossings of the road link (the window principle). Ideally, the time provided by the window will exceed the closure timing required for the crossing, but

Figure 3. Schematized section of real-time-space diagram for afternoon eastbound traffic at railway crossing C.

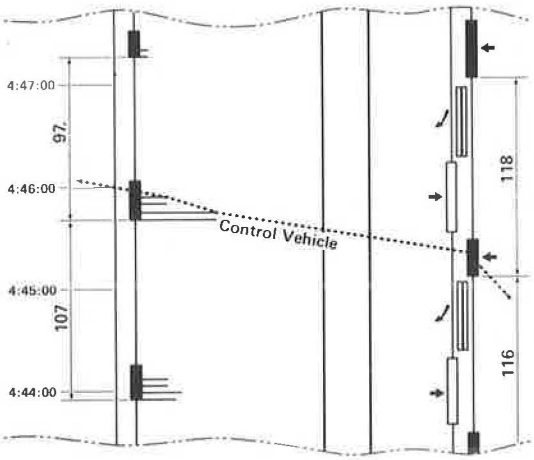


Figure 4. Schematized section of real-time-space diagram for morning eastbound traffic at railway crossing C.

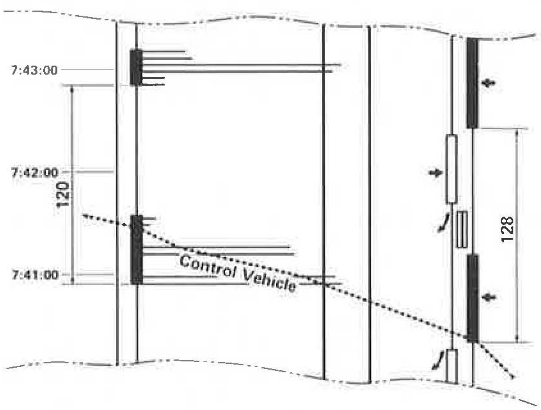


Figure 5. Light rail transit schedule and induced road closures (example).

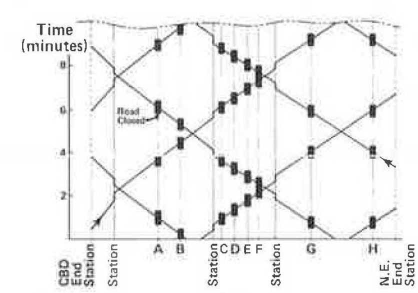
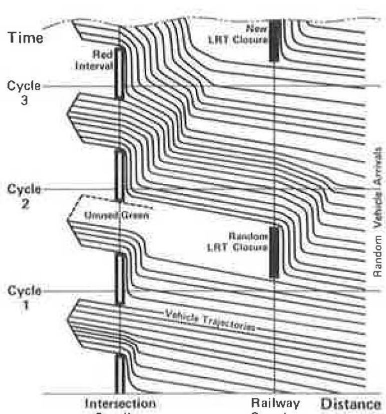


Figure 6. Schematic illustration of capacity losses, increased delays, increased number of stops, and extensive queuing caused by light rail transit crossing closure.



this is a difficult task because of the number of other constraints, such as the LRT scheduling and operation.

The following measures will be used to integrate the intersection control and the LRT operation:

1. The fixed-signal cycle lengths will be defined as an integral fraction of the LRT headways.
2. The LRT will be scheduled to arrive at the crossings during periods protected by red signals at adjacent intersections.
3. In the critical crossing area (area C), the traffic-control system will send a stop signal to the adjacent LRT station. This signal will be programmed so that, when it is released, trains will leave the station to reach the crossing at a time when a window is available. The signal will be transmitted once in shorter cycle lengths and more frequently during longer cycle lengths.
4. The operating speed of the LRT will be influenced by the traffic-control requirements. The goal will be to pass the trains through the crossings without stopping.

Figure 7. Comparison of crossing-control alternatives.

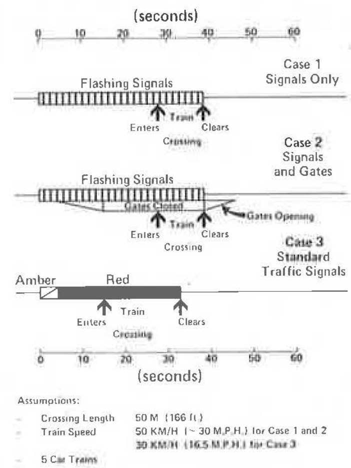


Figure 8. Locations of light rail transit crossing-control and detection equipment.

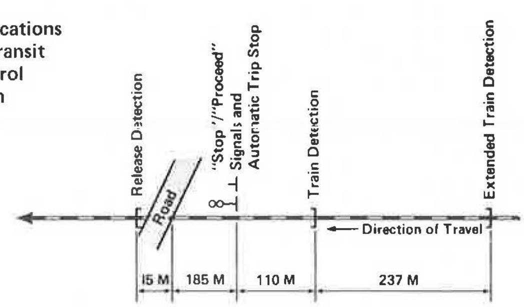
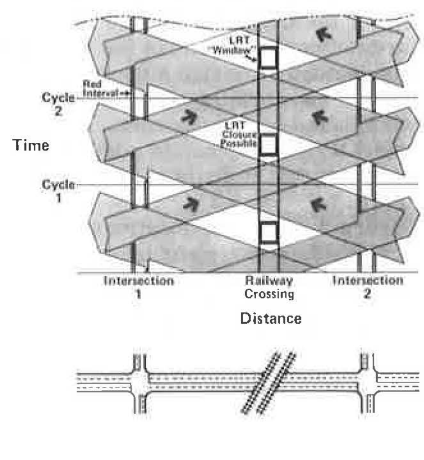


Figure 9. Schematic illustration of the window principle.



5. Although the railway train controls cannot be integrated with the traffic-control measures to the same degree as can the LRT controls, the crossing signals and gates will operate in concert for both modes.

The third principle is the incorporation of special features that are required to guarantee safety within the constraints imposed by the principles of coordination and integration. The major goal is the prevention of queues caused by railway trains, accidents, construction works at adjacent sections of the road network, failures of the control equipment, disruptions in the LRT operation, or the frequent special events in the adjacent recreational facilities. The control algorithm for these special circumstances is based on the detection of unusual queues. The subsequent actions that occur are

1. Warning drivers of the queue or the blockage ahead and advising them to keep the track area clear,
2. Restraining the traffic inflow at upstream intersections,
3. Preferentially treating phases that can relieve the congestion on critical roadway links, and
4. Introducing special phase sequencing that will maintain traffic flow in the directions unaffected by the crossing closure.

These special features will be used individually or in combinations and may be especially useful in areas where the LRT schedule adherence is questionable and where the self-recovery and window principles will be difficult to implement.

EQUIPMENT REQUIREMENTS

The hardware requirements are based on functional principles and design. They must, however, be somewhat flexible to accommodate changes in control tactics and traffic patterns. The basic equipment functions are as follows:

1. Each group of intersections adjacent to the LRT crossings will operate as a traffic-control zone that is characterized by coordinated fixed-time operation and the availability of five independent signal programs. An independent signal program is defined as one having unconstrained choice of the following: cycle length, offsets, interval sequence (program structure), and interval timing.
2. Traffic-control zones containing LRT crossings will also be coordinated in real time. The reference timing (time datum) will be reestablished at regular intervals, despite the fact that individual zones will operate with different cycle lengths. At the beginning, the program changes will be initiated by a time switch.
3. Special features will be implemented at the traffic-control zone level. They will use standard signal-control measures, similar to force off, hold, and skip phase (interval).
4. Special features will not disturb the background control program, i.e., the system will have the capability to restore fully coordinated operation immediately after traffic conditions return to normal.
5. Within the limitations described above, the system will respond to special demands (such as queue detection).

FUTURE CONSIDERATION

Some of the most critical LRT crossings will be replaced by new grade separations in the future. However, in addition to the crossings that will be retained, new grade crossings may be introduced as the LRT network is extended.

The system designed for the first line will be automatically monitored by using the available LRT control hardware. This operating experience will be an important input in the design of special features of the Edmonton computerized transportation-management system.

The use of railway rights-of-way for LRT corridors may be attractive in other cities also, and similar problems with grade crossings may be encountered. These problems should be considered early in the planning process.

ACKNOWLEDGMENT

We wish to gratefully acknowledge the cooperation of the Canadian Transport Commission, the Edmonton Rapid Transit Project, the Edmonton Transit System, and the Edmonton Traffic Operation Section, who all significantly contributed to this project.

REFERENCES

1. Board of Transport Commissioners for Canada. General Order No. E-6.
2. Board of Transport Commissioners for Canada. General Order No. E-13.
3. D. I. Robertson. Cyclic Flow Profiles. Traffic Engineering and Control, Vol. 15, June 1974, pp. 640-641.
4. J. Moe. Train-Activated Rail-Highway Protection. Proc., National Conference on Railroad-Highway Grade-Crossing Safety, Ohio State Univ., Aug. 29-31, 1972.
5. National Electric Manufacturers Association. Proposed Standards Publication for Traffic Control Systems.
6. S. Teply. Coordinated Traffic-Actuated Signals (in German). Strassenverkehrstechnik, Vol. 15, May-June 1971, pp. 88-93.
7. W. H. Morris, Jr. Comparison of Busway and Light Rail Modes. TRB, Special Rept. 161, 1975, pp. 50-61.
8. R. D. Touton, Jr. Electrification and Control Systems for Light Rail Systems. TRB, Special Rept. 161, 1975, pp. 86-92.
9. G. B. Beetle. Light Rail Transit Construction Costs. TRB, Special Rept. 161, 1975, pp. 115-121.

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

Development and Application of a Railroad-Highway Accident-Prediction Equation

Robert A. Lavette, Florida Department of Transportation

This paper reports the development of an accident-prediction equation for train-vehicle collisions at railroad-highway grade crossings that can be used as the basis for the establishment of a priority order for signal improvements. Most of the quantitative and physical factors in the grade-crossing environment were included. Of the 6000 public grade crossings in Florida, 1140 on state roads were used as the study base. The accident-prediction model was developed by the use of a stepwise regression analysis and three unconventional statistical techniques: (a) the analysis of the plots of the residuals, which indicated that a transformation was required (with the transformation of the dependent variable to a logarithmic form, the plot of the residuals was reasonably symmetric); (b) the observed interaction between the independent variables, which resulted in the use of dummy variables, particularly those for active (warning devices) times daily traffic and number of trains; and (c) a bias in the accident prediction that was introduced by the use of logarithms and eliminated by use of a nonlinear least squares adjustment. The accident-prediction model had a multiple correlation of 0.43. The independent variables in the model were the traffic, number of trains, vehicle speeds, train speeds, number of lanes, and presence of warning devices. The accuracy of the accident-prediction equation was demonstrated by comparisons of actual accidents to predicted accidents. The actual number of train-vehicle accidents in 1975 was 70 percent of the number predicted by the model. In 1975, the total number of accidents remained unchanged from that in 1974, but the number of train-vehicle accidents decreased 22 percent.

This paper presents a method for developing an accident-prediction equation for train-vehicle collisions at railroad-highway grade crossings and illustrates the benefits of such an equation to a transportation agency responsible for establishing a priority order for signal improvements. A stepwise regression analysis and three statistical techniques (transformation of data, use of dummy variables, and transformation of the accident-prediction model to its original scale) not previously employed in the development of an accident-prediction model were used.

In July 1972, the Florida Legislature authorized the Florida Department of Transportation (FDOT) to

> determine and adopt a program . . . for the construction cost of projects for the elimination of hazards of rail-highway crossings. Every railroad company . . . shall, upon reasonable demand and notice from the DOT, install, maintain, and operate at such a crossing an automatic flashing light signal, the design of which shall be approved by the FDOT.

The grade crossings on state-maintained roads had been previously inventoried. After the 1972 legislation, the remaining 5000 public grade crossings on local streets were also surveyed. The bases for the data collected were those physical factors that influence accidents (2). The field survey, which was conducted jointly by the FDOT and the railroad companies in 1973, was also part of the Federal Railroad Administration-Association of American Railroads national crossing survey. (This survey confirmed the earlier survey of grade crossings on state-maintained highways.)

The first step in the development of an accident prediction model for rail-highway grade crossings was a review of the existing statistical data and the previous publications on the subject and led to the following conclusions (1):

1. It is imperative that the data base be as accurate as possible. Previous studies did not indicate its verification.
2. Previous reports did not take advantage of all of the statistical techniques available, such as analysis of the residuals and dummy variables.
3. The theory of linear statistical models could be applied conveniently to the available data.
4. Many variables involved in a train-vehicle accident were not included in the data; thus, any model selected would have considerable inherent variation in the number of accidents at a particular crossing.

STATISTICAL ANALYSIS OF THE DATA

Data

The main object was the determination of the relative influences of selected physical factors on the number of train-vehicle collisions at rail-highway grade crossings in Florida. The data analyzed were limited to those features that FDOT could modify and for which complete data were available. For example, there were no data available for the analysis of driver behavior, the optical effectiveness of the railroad signals, or driver-traffic characteristics. A complete listing of the data collected and analyzed is shown below.

1. Maximum posted train speed in miles per hour,
2. Average number of trains per day,
3. Highway system (e.g., Federal-Aid Primary),
4. Rural or urban location of crossing,
5. Road facility (e.g., arterial),
6. Number of lanes,
7. Posted crossings speed limit in miles per hour,
8. Average daily traffic in units of 1000 vehicles/d,
9. Warning device, type 1 (crossbucks, flashing lights, and such),
10. Warning device, type 2 (illumination or stop sign),
11. Minimum approach distance in feet (sight distance to crossing),
12. Parallel road characteristics (within 61 m of track),
13. Minimum clear sight distance in feet (triangle to train),
14. Quadrants with minimum clear sight distance,
15. Maximum clear sight distance in feet,
16. Quadrants with maximum clear sight distance (I, II, III, or IV),
17. Year most recent protection device was installed,
18. Rate of accidents for 5-year period,
19. Number of accidents in 1967,
20. Number of accidents in 1968,
21. Number of accidents in 1969,
22. Number of accidents in 1970,
23. Number of accidents in 1971, and
24. Total number of accidents for 5-year period.

(The model derived in this paper is designed for U.S. customary units; therefore the variables in the tables and equations are not given in SI units.)

The data sources included the FDOT annual inventory and traffic counts (average annual daily traffic) of all state-maintained roads and the field inventory, which confirmed the physical data, including the measurement of approach and triangle (quadrant) sight distances. The train speeds and the number of trains per day were obtained from railroad company timetables and verified by station masters.

Since a careful review showed that the accident history for 1967 (25 percent fewer accidents than in other years) was unreliable, only the data for the years 1968 through 1971 were used in the study. Of the 1155 state-maintained crossings, 1140 were selected for use. Although there were data available for 225 city-maintained crossings, these were on one railroad line with the same number of trains per day, and with no variation in an important independent variable, its regression coefficient could not be determined. Therefore, that sample was not further investigated.

Tables of the basic characteristics of the state-maintained crossings were compiled. Good distribution was obtained for train speed, vehicle speed, traffic counts, and land use (urban, rural, or municipality with a population of less than 5000 persons). The number of trains per day was predominantly in the lower range: Of 1140 crossings, 622 had fewer than 5 trains/d and 1048 had fewer than 15 trains/d. Only 10 crossings had more than 30 000 vehicles/d. (The FDOT is not confident of the accident predictions for crossings with traffic counts above 30 000 and those with fewer than 1 train/d, since the predictions appeared too high.)

Basic Model

Only those statistical methods not discussed in the previous report (1) are fully discussed in this report.

Analysis of the Residuals

The multiple regression model shown below implies that the relation between the dependent variable (y_j) and the ith dependent variable is linear (1).

$$y_j = \beta_0 + \beta_1 x_{1j} + \ldots + \beta_k x_{kj} + \epsilon_j \quad (j = 1, \ldots, N) \tag{1}$$

where

- y_j = observed number of accidents at crossing j for a particular time period,
- x_{kj} = value of kth characteristic for crossing j (assumed constant for the specified time period),
- β_k = regression coefficient for kth crossing characteristic, and
- ϵ_j = unexplained residual variation.

The usual statistical assumptions for a regression model are that the errors ϵ_j have a mean of zero and a constant variance, and are uncorrelated and normally distributed. These assumptions are never exactly satisfied in real life.

Again, if y_j, the independent variable, is the number of accidents per year, the variance of y_j will likely depend on the level of the x_{ij}, i.e., in effect on $\sum_{i=1}^{k} \beta_i x_{ij}$. The reason for this is, in part, that the dependent variable must have at least Poisson variation, and the variance of the Poisson distribution is equal to the mean. To stabilize the variance, a transformation is needed and the type of transformation will be suggested by an examination of the residuals. . . .The original form of the dependent variable y_j was the raw number of accidents for the period observed at crossing j. On the basis of the Poisson nature of accidents at a crossing, regressions involving the raw score (y_j) were quickly discarded. Plots of the residuals ($y_j - \hat{y}_j$) against the independent variables such as the average number of trains per day, indicated that the distribution was highly skewed and some transformation was in order.

The same was true for independent variables such as train speed and average daily traffic.

Data Transformations

A square-root transformation was attempted prior to the logarithmic transformation since this would stabilize the residual variance at 0.25 for a Poisson random variable, but the plot of the residuals ($y_j^{1/2} - \hat{y}_j^{1/2}$) against the average daily traffic unfortunately exhibited considerable skewness (1).

The square-root model produced a higher multiple correlation than the model finally adopted (R = 0.4528 against R = 0.4285 respectively), but the increase was not judged to outweigh the advantage of the symmetric distribution of residuals produced by the logarithmic model.

There may be some question as to why none of the regressions reported here explained more than 20 percent (100 R^2) of the variance while those reported by Coleman and Stewart (8) explained from 63 to 78 percent. This may be due to the fact that, besides the differences in sample size and area covered, Coleman and Stewart were fitting to grouped data so that their R^2 figures pertain to variances of group means. The variance of individual crossings within groups is not considered, although it would have a substantial effect on the variability of a prediction for a single crossing.

From the report by van Belle, Meeter, and Farr (1),

The residual mean square, after fitting the square-root model, was 0.288 as compared to the theoretical variance of 0.25. Since the variance about the mean (for the square-root model) was 0.359, a significant reduction in variance was produced by fitting the model, but a substantial amount of variation (0.288 versus 0.250) remained unexplained.

Logarithmic transformations on independent variables that had large coefficients of variation were also positively skewed.

The transformation $\ln(y_j + a)$ was selected with the value of a determined from a plot of the residuals. Initial values of a tested were 1, ½, and . . . finally 0.04.

The residuals were reasonably symmetric for a = 0.04. An additional reason for this particular choice of a is that an analysis on an annual basis using $\ln[(y_j/4) + 0.01]$ would differ from the present analysis by only a constant.

Partitioning the Sample

Besides the investigation of several transformations, the data were partitioned into two samples by the rural-urban as well as the active-passive dichotomies. This approach resulted in equations which fitted the data as closely (in the sense of multiple R) as the approach finally adopted: i.e., the dummy variable techniques. The latter was used because dummy variables allow the selective interaction of variables.

However, partitioning the sample produces separate estimates for all of the parameters in both samples, whether significantly different or not. Also,

dummy variables allow the examination of more combinations of dichotomies without reducing the sample size.

Finally, the splitting of the sample on the basis of the active-passive warning devices created the problem of crossings whose warning devices were improved during the study period. These crossings had to either be discarded or be allocated into both the active and passive groups, which would artificially increase the sample size.

Influence of Sight Distances

One aspect of the partitioned samples that was not present in the final model is the introduction of independent variables of sight distance in relation to the required stopping sight distance. These two variables were the ratio of the available approach sight distance to the crossing to the desirable required stopping sight distance and the ratio of the quadrant clear sight distance (the sight triangle of the approaching train measured along the road) to the desirable required stopping sight distance (3). The quadrant clear sight distance is similar to that described by Schoppert and Hoyd (2) and is the distance from the tracks at which a line of sight to the approaching train would be obstructed.

The data were partitioned into rural versus urban area crossings and also into active versus passive warning-device crossings to observe the effect of sight distance. Those crossings at which the warning devices were modified during the study period were eliminated from the active versus passive partition. The approach sight distance variable was significant only for the urban partition with an F-value of 3.8 (multiple R was 0.46). In Florida, with its flat terrain, there are very few crossings with restricted approach sight distances, particularly if the stopping distances for dry pavement are used. The required stopping sight distance is based on a wet pavement condition.

The effect of the clear sight distance variable on the partitioned models was significant. The F-values for the rural and passive partitions were 7.9 and 5.7 respectively (the multiple Rs were 0.40 and 0.51 respectively). The critical F-value at the 0.05 significance level was 3.84, whereas the F-values for the urban and active partitions were only 0.2 and 1.4 respectively. The low highway speeds in urban areas and the fact that most crossings in these areas have active (train-activated) warning devices mean that sight distances have minimum significance in urban areas or at crossings with active warning devices. These models did not fully use the analyses of residuals or dummy variables. In retrospect, two mistakes were made in analyzing the sight-distance data. One was that dummy variables should have been used for the active versus passive condition. The other was that, since it was unusual for wet pavement to be a contributing factor in a train-vehicle collision, the required stopping distances for dry pavement should have been used (for study purposes only, not for design practices). Thus, if properly analyzed with the use of dry pavement conditions, the clear sight distance would be a significant variable, but, in Florida, the approach sight distance would not be significant.

Of the 1140 grade crossings surveyed, 95 percent had at least one quadrant in which driver vision was obstructed, whereas the sight distance to the crossing itself was obstructed for only 9 percent of the crossings. These sight distances were calculated and surveyed by using the required stopping distance on wet pavements (plus a perception and reaction time of 2.0 s) (3).

The independent variable used to determine the significance of quadrant sight distance was

$$\text{ratio } C = (\text{minimum clear sight distance}/\text{required stopping sight distance}) \times 10 \qquad (2)$$

For a dry pavement condition, not only is the denominator (the required stopping sight distance) reduced but also the hypotenuse and base of the sight triangle are substantially reduced. The field survey was conducted by observing a point on the track that formed the angle of the base and hypotenuse of the sight triangle. The base distance along the track (the critical approach distance) is the distance a train travels at its maximum allowable speed during the time that it takes a passenger vehicle to stop. Thus, if the passenger-vehicle stopping time were reduced, the base distance along the track would also be reduced. Since the minimum clear sight distance is the distance along the highway from the track to a point where the driver cannot view a train at its critical approach distance, the number of crossings where the minimum clear sight distance equals the required stopping sight distance would be increased.

Selection of Dummy Independent Variables

To account for the effects of automatic train-warning devices on the other independent variables, categories of basic warning devices were established as shown below.

	Variable			
Category	Flashing Lights (PD211)	Gates (PD27)	Advance Light (PD29)	Active-Passive Code
Passive	0	0	0	A = 0
Active 1 (flashing lights)	1	0	0	A = 1
Active 7 (flashing lights and gates)	1	1	0	A = 1
Active 9 (A7 plus traffic signal preempted)	1	0	1	A = 1

Thus, PD211 (flashing lights) is nonzero only when there are active warning devices; PD27 (flashing lights and gates) is nonzero only when there is gate protection. The regression coefficients associated with these variables indicate the additional reduction in ln(number of accidents + 0.04) when a particular automatic warning device is present (1). For example, the regression coefficient associated with PD211 estimated the additional reduction in the dependent variable due to an active warning device. All three variables were included in a regression equation with coefficients b_1, b_7, and b_9 if active devices 1, 7, and 9 were present at this crossing. Thus, a priori, the coefficients are expected to be negative (warning devices should decrease the number of accidents), and b_7 should be greater than b_9 since the most protection should produce the greatest decrease.

One problem associated with the data was that the warning devices at some of the crossings were modified during the study period. The assumption was made that... such modifications occurred at the midpoint of the year and the nominally 0-1 variables for these crossings were coded as follows:

	Year of Modification				
	1968	1969	1970	1971	
1(Active)	$\frac{7}{8}$	$\frac{5}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	0(Passive)

Thus, if a crossing were modified from passive (dummy code 0) to active (code 1) in the year 1970, it was assumed to be active for 3/8 of the 4-year study and was coded with this value.

Similar dummy variables were established for the rural versus urban categories. The 25 dummy variables listed in the table below were derived from the basic variables listed in Table 1.

Variable	Description	Variable	Description	Variable	Description
15	A × 1	24	R × 1	32	R × 13
16	A × 2	25	R × 2	33	R × A × 1
17	A × 6	26	R × 4	34	R × A × 2
18	A × 7	27	R × 5	35	R × A × 9
19	A × 8	28	R × 9	36	R × A × 10
20	A × 9	29	R × 10	37	R × A × 11
21	A × 10	30	R × 11	38	R × A × 12
22	A × 11	31	R × 12	39	R × A × 13
23	A × 12				

During the development of the regression model, considerable attention was given to the interactions among the independent variables.

Variables can interact, that is, their joint effect on the dependent variable could be markedly different from the sum of their individual effects. For example, the effect on the accident rate of adding active warning devices . . . varies as the average daily traffic varies. To allow for this, additional independent variables were constructed by multiplying the active-passive dummy variable A (flashing lights and flashing lights and gates) and the rural-urban dummy variable R by other independent variables. These variables are denoted by A × 1, A × 2, . . . , R × 1, R × 2, etc.

For example, A × 1 is the interaction of a kind of automatic warning device with variable 1 (ln of maximum posted speed). The actual selection of the variables and the interactions entered into the program was often the result of knowledge of the grade-crossing environment and not necessarily the result of the stepwise regression procedure. Some variables, such as the crossing speed limit, can be altered, but others, e.g., the location of the crossing, cannot. In this kind of situation, it is not helpful to say that urbanization causes more accidents at grade crossings.

Although certain interaction variables were forced into the regression program to observe their effect, they did not improve the final model. For example, when the variable A × 1 (ln of the posted maximum train speed) was forced into the model, it had an acceptable F-value, but another variable, A × 2 (ln of the number of trains per day) dropped out.

Accident-Prediction Model

The final stepwise regression analysis, after 20 analyses, involved the 39 independent variables listed above and in Table 1. The standard error of estimate for the final regression was 1.52 and the multiple correlation was 0.43. Eight independent variables, shown in Table 2, had F-ratios greater than 7. The critical F-value at the 0.05 significance level is 3.84; however, the increase in predictive variance precluded the addition of other predictor (independent) variables. The model selected was

$$
\begin{aligned}
\text{predicted } \ln(y + 0.04) = & -8.0757 + 0.4368\,[\ln(\text{ADT})] \\
& - 0.1440[\text{A} \times \ln(\text{ADT})] \\
& + 0.3178[\ln(\text{maximum train speed})] \\
& + 0.4838[\ln(\text{number of trains per day})] \\
& - 0.3180[\text{A} \times \ln(\text{number of trains per day})] \\
& + 0.3870[\ln(\text{crossing speed limit})] \\
& + 0.2249(\text{A} \times \text{number of lanes}) \\
& - 0.4662(\text{PD27})
\end{aligned} \tag{3}
$$

where y is the total number of accidents for 1968-1971 (1).

Four of the eight variables are expected to be involved in any model for accident prediction at grade crossings: daily traffic volume, maximum train speed, number of trains per day and crossing speed limit. All of these independent variables have positive regression coefficients.

Hence, they are positively correlated with the number of accidents at a grade crossing.

The other four independent variables involve the nature of the warning device at a crossing. The first of these A × ln (ADT), i.e., the effect of an active rather than a passive warning device, varies with the level of the traffic volume. In particular, for crossings with passive signing, the predicted ln(y + 0.04) is increased by 0.4368 for each unit increase in ln(ADT), whereas for crossings with active warning devices, the predicted ln(y + 0.04) is increased by (0.4368 - 0.1440) for each unit increase in ln(ADT). This is because the variable [A × ln(ADT)] is nonzero only for crossings with active warning devices. The interpretations for the other interactions are similar.

Some of the active versus passive dummy variables are highly correlated, such as A × 10 (crossing speed limit) and A × 2 (number of trains per day) or A × 9 (number of lanes) and could have been substituted for each other with little effect on the predictive accuracy of the fitted equation.

Transforming the Accident Prediction

The use of the logarithmic transformation for the dependent and independent variables gives a statistical model that more closely satisfies the least squares assumption. This model also provides a method for scheduling crossing improvements that is based on accident prediction on a logarithmic prediction scale. When the logarithmic form of the model (y = 0.04) was transformed to the original scale, a substantial negative bias was introduced. To obtain an unbiased transformation, Beauchamp and Olsen (5) used a complicated procedure to derive estimates for the mean of a lognormal variable that depends on a single independent variable, but a simpler approach was used here. The objectives were that the sum of the predicted accident rates equal the sum of the actual number of accidents; that all predictions be nonnegative; and, subject to this, that the predictions should satisfy a least squares property. Thus, for $i = 1, \ldots, 1140$, x_i = 4-year total accidents at crossing i, $\hat{y}_i$ = least squares estimate of rate obtained from crossing i [in the ln(x + 0.04) scale], and $\hat{x}_i$ = predicted x_i in the original scale.

Let

$$\hat{x}_i = \exp(\alpha\beta\hat{y}_i) \tag{4}$$

where we should obtain $\Sigma x_i = \Sigma x_i$. Let $T = \Sigma x_i$; then

$$\exp(\alpha) = T/\Sigma(\exp\beta\hat{y}_i) \tag{5}$$

so that the estimator now depends only on β, i.e.,

$$\hat{x}_i = [T/\Sigma(\exp\beta\hat{y}_i)]\,(\exp\beta\hat{y}_i) \tag{6}$$

A value of β is chosen to minimize S where

$$S = \left\{x_i - [T/\Sigma(\exp\beta\hat{y}_i)]\,(\exp\beta\hat{y}_i)\right\}^2 \tag{7}$$

A computer program written to evaluate S gave values of $\alpha = 1.109$ and $\beta = 0.968$ (1).

Table 1. First-order variables used in stepwise regression analysis of accidents at crossings on state-maintained roads.

Variable	Description	Mean	Standard Deviation
1	ln of posted maximum train speed[a]	3.373	0.673
2	ln of number of trains per day[a]	1.482	0.891
3	PD211 [flashing lights (warning device)]	0.504	0.470
4	PD27 (flashing lights and gates)	0.127	0.323
5	PD29 (flashing lights, gates, and preemption)	0.007	0.081
6	Rural versus urban, category 2, small municipality	0.161	0.386
7	Rural versus urban, category 3, urban characteristics	0.072	0.258
8	Rural versus urban, category 1, rural	0.503	0.500
9	Number of lanes	2.493	0.993
10	ln of crossing speed limit[a]	3.698	0.332
11	ln of average daily traffic[a]	7.715	1.503
12	Ratio C[b]	2.733	2.645
13	Ratio D[c]	9.592	1.450
14	ln of (total accidents - 0.04)[a]	-2.120	1.677

[a] Logarithms are to base e.
[b] Ratio C = coefficient of impaired view of approaching train = 10 × minimum clear sight distance/required stopping sight distance.
[c] Ratio D = coefficient of impaired view of a crossing protection device = 10 × minimum approach distance/required stopping sight distance.

Table 2. Final stepwise regression analysis: variables retained, regression coefficients, and F-values (state-maintained roads only).

Step	Variable	Entered Description	Final Regression Coefficient	F-Value	R-Correlation
1	11	ln(avg daily traffic)[a]	0.436 8	139.1	0.280 8
2	22	A(active) × 11 (avg daily traffic)	-0.144 0	16.9	0.343 0
3	1	ln(maximum train speed)[a]	0.317 8	15.6	0.379 5
4	2	ln(number of trains per day)[a]	0.483 8	30.1	0.395 1
5	16	A × 2(number of trains per day)	-0.318 0	7.3	0.406 9
6	10	ln(highway speed limit)[a]	0.387 0	7.8	0.415 7
7	20	A(active) × 9 (A × number of lanes)	0.224 9	9.0	0.421 8
8	4	PD27 (flashing lights and gates)	-0.466 2	7.9	0.428 5
—	—	intercept (b_0)	-8.075 71	—	—

Notes: Dependent variable is ln(number of accidents + 0.04); multiple correlation R = 0.4285; standard error of estimate = 1.52; F-ratio = 31.793; and critical F-value = 3.84 at 0.05 confidence level.

[a] Logarithms are to base e.

MODIFICATION OF MODEL

Modification of Regression Coefficients

The regression model selected indicated that when the coefficient for the number of lanes was entered into the model (0.225 × number of lanes) its effect offset the value of the gates (-0.466). From an engineering standpoint, it is obvious that gates of adequate length will drastically reduce the sight restrictions at multilane roads, where a driver's view of signal lights, for example, could be obstructed by a truck in an adjacent lane. However, two-lane roads are the normal condition and should not increase the risk when active warning devices are present. Thus, the number of lanes was allowed to affect the model only when there were more than two lanes and to not affect the model when gates were present. To offset this change, the coefficient for gates was reduced from -0.466 to -0.233.

Sight distances—the ability of the driver to view the approaching train and to see the warning signs or flashing lights—definitely are part of train-vehicle accident prediction and consequently are part of the accident-prevention environment, and these independent variables were significant when the data were partitioned into rural versus urban or active versus passive categories.

The coefficients for C (clear sight distance) and D (approach sight distance) in the stepwise regression models apply only to passive signing for C, and only to crossings in urban areas for D. These coefficients were used as a guide to derive the following terms:

$$C = 0.33 - (\text{minimum clear sight distance}/\text{required stopping sight distance}) \times 10 \times 0.123 \quad (8)$$

When C is less than zero, this term is not used.

$$D = 0.28 - (\text{minimum approach distance}/\text{required-stopping sight distance}) \times 10 \times 0.028 \quad (9)$$

Examination of the clear sight distance term shows that it increases the accident prediction only when the distance from the track at which the driver can first view an approaching train is no more than one-fourth of the required stopping sight distance (3) on wet pavement. Of the 1140 grade crossings examined, 65 percent had at least one quadrant in which this occurred. However, the minimum approach distance term increases the accident prediction whenever the approach (sight) distance (the ability to see the crossing) is less than the required stopping sight distance. But, only 9 percent of the crossings had any sight-distance restriction to the crossing.

Calculations of the reduced clear sight-distance triangle made by using the stopping distance for dry pavement and assuming that the obstacle restricting the view of the approaching train was 4.6 m (15 ft) from the edge of the travelway showed that, if the minimum clear sight distance were three-fourths of the required stopping sight distance under wet pavement conditions, then there would be sufficient sight distance available for dry pavement conditions. Therefore, of the 1140 crossings examined, 60 had adequate clear sight distance on wet pavements and 125 had adequate clear sight distance on dry pavements.

The variables for restricted approach sight distances (D) and the restricted clear (triangle) sight distances (C) were included so that, when the actual regression coefficients were obtained at a later date, new terms would not have to be added to the existing computer programs. The final model used by FDOT later proved to be satisfactory, and the planned subsequent regression analysis was not undertaken.

Accident-Prediction Equations

By using the final stepwise regression model and the modifications discussed below, two equations were established. The first (Equation 10) calculates the accident potential (t_p) for 4 years at grade crossings with only passive signing. The second (Equation 11) calculates the accident potential (t_a) for grade crossings with active warning devices.

$$\begin{aligned} t_p = & -8.075 + 0.318 \ln S_t + 0.484 \ln T + 0.437 \ln A \\ & + 0.387 \ln V_v + [0.28 - 0.28(\text{MASD}/\text{RSSD})] \\ & + [0.33 - 1.23(\text{MCSD}/\text{RSSD})] \\ & + 0.15 \quad \text{(if no crossbucks)} \end{aligned} \quad (10)$$

$$y = [\exp(0.968 t_p + 1.109)]/4 \quad (10a)$$

$$\begin{aligned} t_a = & -8.075 + 0.318 \ln S_t + 0.166 \ln T + 0.293 \ln A \\ & + 0.387 \ln V_v + [0.28 - 0.28(\text{MASD}/\text{RSSD})] \\ & + 0.225(L - 2) - 0.233 \quad \text{(if gates)} \end{aligned} \quad (11)$$

$$y = [\exp(0.968 t_a + 1.109)]/4 \quad (11a)$$

where

- A = vehicles per day or annual average daily traffic,
- L = number of lanes,
- MASD = actual minimum stopping sight distance along roadway,
- MCSD = clear sight distance (ability to see approaching train along the roadway, recorded for the four quadrants established by the intersection of the railroad tracks and road),
- RSSD = required stopping sight distance on wet pavement,
- S_t = maximum speed of the train,
- T = yearly average of the number of trains per day,
- t_a = ln of the predicted number of accidents in the 4-year period at crossings with active protection,
- t_p = ln of the predicted number of accidents in the 4-year period at crossings with passive protection,
- V_v = posted vehicle speed limit unless geometrics dictate a lower speed, and
- y = predicted number of accidents per year at crossing.

(The variable [0.33 - 1.23(MCSD/RSSD)] is omitted if the crossing is protected by flagmen or the calculation is less than zero, the variable [0.28 - 0.28(MASD/RSSD)] is omitted if sight restriction is due to a parallel road, and the variable (L - 2) is omitted when there are gates.)

Adjustment for Accident History

The stepwise regression model is a reasonable accident predictor for each grade crossing, which admittedly would be biased by the introduction of the accident history of a crossing. It is also possible that the phenomenon of regression toward the mean may mean that a crossing that has two or three accidents in 1 year may not have any more until it reaches its actual predicted accident rate. However, the accident history can be used as an adjustment to compensate for some of the failings of the accident predictor. The need for an accident-history adjustment was based on the following:

1. The present stepwise regression model explains only 18 percent of the accidents that occur (the multiple correlation R was 0.43) because human failure is involved in over 90 percent of them. Although it would be possible to increase the multiple correlation by taking into account different driver profiles at various crossings, it would be impossible to collect such data.
2. Accident histories are used by engineers to identify deficient systems, and in the event of a lawsuit, it would be difficult to explain in court why the accident history at a particular location had been ignored.

Thus, an accident-history adjustment equation that would increase but never decrease the accident predictor was used. This adjustment for accident history is calculated only when the accident history is greater than the accident prediction.

$$y' = y(H/yP)^{1/2} \quad (12)$$

where

- y = accident prediction based on the regression model,
- y′ = accident prediction adjusted for accident history,
- H = number of accidents for a 6-year history or since the year of the last improvement, and
- P = number of years of the accident-history period.

The accident-history adjustment has not been a major factor in determining the most hazardous grade crossings. Of the 98 crossings with the highest accident prediction, 61 were not affected by the accident history. The use of accident history has, however, helped to identify grade crossings with unique problems that were not identified in the accident-prediction model.

USE OF ACCIDENT-PREDICTION EQUATIONS

Selection of Grade-Crossing Improvement Projects

A simple method of rating each grade crossing from 0 to 90 was derived from the accident-prediction model. This method, entitled the safety index, was used to rank each grade crossing. A safety index of 70 is considered safe (no further improvement is necessary): A grade crossing with an accident prediction of 0.05 or one accident every 20 years would have a safety index of 70. It is not economical to provide active warning devices at grade crossings having lower accident-prediction indexes. A safety index of 60, or one accident every 9 years, would be considered marginal.

Each grade crossing is assigned a statewide priority number based on the safety index, i.e., the grade crossing with the lowest safety index would be assigned priority one. If there were no fund limitations, the selection of grade crossings for an improvement program would be simplified. However, the funds for the program, which are received primarily from the 1973 and 1976 Highway Safety Acts, are divided between Federal-Aid routes and off-system routes. Since these funds have become available, FDOT has scheduled 125 grade crossings on Federal-Aid routes for improvement at a cost of $5.8 million. As of June 1, 1977, 90 of them had been completed. When the total 125 are completed (this does not include the 330 urban streets that were recently added to the Federal-Aid system), all grade crossings on Federal-Aid routes that had a safety index of less than 70 will have automatic warning devices.

However, there will still be the major problem of the 4460 grade crossings on off-system routes. At the beginning of the program, 4150 of these had only passive warning signs; 400 did not even have crossbucks. FDOT has scheduled 130 of them for improvement, and as of June 1, 1977, 80 had been completed.

Reduction of High-Accident Sections

The regression analysis showed that reducing train or vehicle speeds or both reduces the probability of accidents. Also, since the accident prediction increases the logarithm of the number of vehicles, the same number of vehicles using fewer grade crossings will reduce the probability of accidents. Thus, the closing of any grade crossing will decrease the accident probability; i.e., the accident prediction for the one crossing with the combined traffic will be less than the sum of the accident prediction for the two crossings. Of course, if the traffic from a closed crossing is diverted to a grade crossing with automatic warning devices, the probability of an accident will be reduced even more.

A computer program that compares the sum of the accident predictions on each track to a statewide average for a particular category of track was developed. Each combination of urban versus rural and type of train is a different category. For example, if the track is in a small municipality and both freight and passenger trains use it, the category will be town and passenger. The formula used to select a high-accident (abnormal track section is

$$\lambda p = \lambda c + K(\lambda c/T)^{1/2} - 1/2T \tag{13}$$

where

λp = critical accident potential per mile of track,
λc = average accident potential per mile for the category of track being tested,
T = natural logarithm of the average number of trains per day, and
K = constant.

The magnitude of K determines the level of statistical significance and controls the number of track sections that should be investigated. The K-value is 2.567, which means that the probability of the accident prediction on the track section selected being abnormal is 99 percent.

Any section of track where the sum of the accident predictions is higher than λp should be examined to reduce the accident prediction. Any of the following actions can be taken: (a) close unnecessary crossings, (b) install automatic warning devices, (c) reduce train speeds, (d) reduce highway speeds, and (e) construct grade separations. The action(s) taken should depend on the feasibility and benefit-cost studies. The reduction of accidents as one of the benefits is based on the reduced accident predictions for the crossings affected.

Effects of Grade-Crossing Improvements on Accident Reduction

In 1974, Governor Reubin Askew committed FDOT to a massive railroad-highway grade-crossing improvement program. The goal was to reduce fatalities by 50 percent (from 90 to 45) by improving 20 percent (1200) of the 6000 grade crossings in the state in 6 years. The improvement of the grade crossings on the primary highway system had been under way, and this was increased by the use of Emergency Highway Safety Funds. In 1974, after the implementation of the Highway Safety Act of 1973, the rate of grade-crossing improvements was doubled to 120/year, and many grade crossings not on the system were included.

The analysis of the effect of the crossing improvements on the 1974 and 1975 accident rate is complicated by concurrent events, such as the late 1973 oil embargo, the 88.5-km/h (55-mph) speed limit, and a decrease in the 1974 vehicle operating speed. (In 1975, vehicle operating speeds increased but remained below 1973 operating speeds.) Also, the number of vehicle-kilometers traveled varied from 94 800 million in 1973 to 98 000 million in 1974 to 106 000 million in 1975 (53 300 million, 61 900 million, and 62 200 million vehicle-miles traveled respectively), and the number of train movements decreased approximately 10 percent. However, some interesting comparisons still can be made.

Three groups of grade crossings that had no signal improvements in 1974 and 1975 were analyzed. According to the accident-prediction model, these crossings had the highest potential for accidents. The results of the analysis are shown below.

Group	Number in Group	Accident Prediction	1975 Accidents	Number Without Accidents (1969 to 1974)	First Accident in 1975
Highest	98	46.5	34	26	9
Second highest	99	36.9	24	38	5
Third highest	100	25.4	18	35	2

The broad rankings produced by the accident-prediction model are borne out by the 1975 accident experience. For the three groups, the accident experience ranged from 65 to 73 percent (average 70 percent) of the accident prediction. The prediction model is based on the 1968 to 1972 accident (preenergy crisis) data and 1973 vehicle speed limits (maintained for accident predictions), which may be one of the explanations as to why the predictions are higher than the 1975 experience. Only 25 percent of those accident predictions were affected by the accident history.

An examination of the statewide accident trends shows that the number of train-vehicle collisions decreased from 498 in 1974 to 390 in 1975 (22 percent), although the total number of accidents remained unchanged. Among the 108 train-vehicle accident decrease, only 13 can be attributed directly to the installation of automatic warning devices. This leaves a 19 percent decrease in accidents from 1974 that is still unexplained. In 1976, train-vehicle collisions were reduced another 15 percent to 330 (the total number of accidents decreased 4 percent). From 1974 to 1976, the number of fatalities due to train-vehicle collisions decreased from 75 to 55 (25 percent).

These statistics indicate that those crossings having higher accident predictions experienced the most accidents. The accident experience of those crossings without previous accident experience also was proportional to their accident predictions. Also, the accident-prediction model was within 15 to 30 percent of the actual accident experience even with the current downward trend in train-vehicle collisions.

The effect of the installation of automatic warning devices on the number of subsequent accidents was further analyzed. Of those grade crossings modified between July 1, 1974, and October 30, 1975, only 30 were based on the accident-prediction model. (All but 3 of these were installed after July 1, 1975.) However, even these limited results are encouraging. The following results were achieved from the analysis of two groups of 100 grade crossings each.

Crossings	Accident Prediction: Without Modification, One Year	Without Modification, Post Installation	With Modification, Post Installation	1975 Accidents
Modified	24	17	5	4
Unmodified	25	–	–	18

Since the 100 modified grade crossings included those that were modified during 1975, the accident prediction was adjusted downward to include only the time period after the installation of the crossing warning devices. Only those accidents that occurred after the installation were counted. During this period, 5 accidents were expected, but only 4 occurred, which agrees with the control group of unmodified crossings that experienced only 72 percent of their predicted number of accidents.

Thus it appears that the accident-prediction model consistently predicts accidents to within 15 to 30 percent of their actual occurrence, and the reduction in accidents after the installation of automatic warning devices is as expected. The grade crossings that are modified in fiscal year 1976 will provide better data, since three-fourths of them were selected on the basis of the accident-prediction model. During this period, 43 of the 100 most hazardous grade crossings will be modified.

CONCLUSION

The accident-prediction model can be effectively used to develop a grade-crossing improvement program. It identifies groupings of crossings (with or without the accident-history adjustment) that can be expected to experience the most accidents if they are not modified, and the accident experience after modification has been in reasonable agreement with that predicted.

ACKNOWLEDGMENTS

This report could not have been developed without the efforts of G. van Belle, D. Meeter, and W. Farr, who prepared the report, Influencing Factors for Rail-Highway Grade-Crossing Accidents in Florida. In addition, the efforts of Meeter in editing the many rewrites of this report and in providing needed technical assistance are greatly appreciated. Appreciation is also given to Winifred Bailey, who suffered through the typing and editing of the many revisions, and to Wyndal Hand, who assisted with the data.

The changes to the model and the subsequent conclusions as to the sight-distance factors and the effects of gates are those of the author and are not necessarily those of the above-mentioned consultants.

REFERENCES

1. G. van Belle, D. Meeter, and W. Farr. Influencing Factors for Rail-Highway Grade-Crossing Accidents in Florida. Accident Analysis and Prevention, Vol. 7, 1975, pp. 103-112.
2. D. Schoppert and D. Hoyd. Factors Influencing Safety at Highway-Rail Grade Crossings. NCHRP, Rept. 50, 1969.
3. A Policy on Design Standards for Stopping-Sight Distances. American Association of State Highway and Transportation Officials, Sept. 2, 1971.
4. J. B. Hopkins and F. R. Holmstrom. Toward More Effective Grade-Crossing Lights. Transportation Systems Center, U.S. Department of Transportation, Cambridge, Mass., 1975.
5. J. Beauchamp and J. Olsen. Estimates for the Mean and Variance of Lognormal Distribution Where the Mean Is a Function of an Independent Variable. Eastern Deciduous Forest Biome Study, Oak Ridge National Laboratory, Technical Rept. 71-101, 1972.
6. Manual on Uniform Traffic-Control Devices. American Association of State Highway and Transportation Officials and U.S. Department of Transportation, 1971.
7. W. J. Dixon, ed. BMD Biomedical Computer Programs. Univ. of California Press, Berkeley, 1971.
8. J. Coleman and G. Stewart. Accident and Accident-Severity Prediction Equations. Proc., National Conference on Railroad-Highway Crossing Safety, Washington, D.C., 1974.

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

Visual Performance of Drivers During Rainfall

Ron S. Morris, John M. Mounce, Joe W. Button, and Ned E. Walton, Texas Transportation Institute, Texas A&M University

This paper reports an investigation of the effect of rain on the visual performance of drivers. The degradation of static visual acuity in terms of visual angle, detection probability, and legibility as a function of rain intensity was determined by experiments that used a rainfall simulator that produced artificial rain. The significant findings include the following: (a) Water on the windshield is the primary factor accounting for reduced visual performance, (b) visual degradation in the daytime with windshield wipers in operation appears to be a linear function of the rain rate with normal drop sizes, (c) during nighttime conditions, drop size is a significant factor in reducing visual performance (smaller drops are a more serious problem than is the rain rate), (d) wiper speeds above 50 CPM do not improve visual performance, (e) without windshield wipers, visual performance is reduced to levels that are unacceptable for driving (equivalent to visual acuity greater than 20/200) at rain rates greater than 2.5 cm/h (1 in/h), and (f) the effective rain rate can be determined from the vehicle velocity, the terminal velocity of the drop, the rake angle of the windshield, and the actual rain rate.

The factor of visibility during adverse weather has been largely neglected by the highway transportation industry. There are at least two reasons for this: These are that the problems associated with driver visibility have been underestimated and that objective measurements of the effects of wet weather on the visual performance of drivers are difficult to obtain. Thus, there have been very few developments designed specifically to assist the automobile driver in the performance of visual tasks during adverse weather (1).

EQUIPMENT AND METHODOLOGY

The objective tests used in this research determined the effects of selected, controlled intensities of artificial (simulated) rainfall on the visual performance of drivers relative to visual acuity, target detection, recognition, and legibility. These tests were also designed to assess the improvement to driver visibility afforded by windshield wipers at various cyclic rates. All of the tests were conducted on overcast days to more closely simulate actual rain conditions. To eliminate the effects of wind on the paths of the falling drops, the tests were conducted on days and nights when the wind was less than 8 km/h (5 mph) (2). To establish a feasible relation between the simulated condition and real-world conditions, the rainfall characteristics of the stimulus, i.e., the droplet-size distribution, were evaluated and compared to the characteristics of natural rainfall.

The rainfall-simulator studies used controlled, in-vehicle experimentation with human subjects. An overhead pipe and nozzle system (Figure 1) was used to produce rainfall artificially. The simulator was 56.4 m (185 ft) long and had 32 spray bars 7.6 m (25 ft) long.

All of the observations were made by test subjects from a 1975-model automobile in which the windshield, windshield wipers, and headlights were original vehicle equipment that had been maintained at the recommended specifications. The windshield wipers were modified by adding a 430-W (¼-hp) AC gear motor to replace the original motor. An inverter to supply AC power and a variable speed controller to produce any wipe rate from 0 to 80 cycles/min were placed where operation by the test administrator would be convenient.

VISIBILITY PERFORMANCE

Three basic visibility measures were used in the simulator experiments: (a) visual acuity, (b) legibility, and (c) target detection and identification.

Panels containing Landolt rings of nine different sizes (Figure 2), which could be fastened to a portable backboard in any desired orientation, were used as the standard measure of visual acuity. The sizes of the rings were designed to cover the range of visual acuity from 20/20 to 20/200 when observed from a distance of 45.7 m (150 ft). Standard Manual on Uniform Traffic Control Devices (MUTCD) type D1-1 destination signs were observed by the test subjects to determine the legibility distance (Figure 2). These signs also were placed 45.7 m (150 ft) from the point of observation.

Several targets were used in the detection and identification portion of the experiment. These were

1. Mannequin—upright, male caucasian clothed in a dark gray raincoat;
2. Nonreflective sign—small, single-post sign; and
3. Front or rear of object vehicle—light brown, 1967 sedan with no light display.

EXPERIMENTAL DESIGN ADMINISTRATION

The experimental design was structured into two basic tests, S-1 and S-2. The data collection from tests S-1 and S-2 was divided into two sets of observations based on the type of visual performance required of the subject. The first set included the response related to visual acuity and legibility, and the second set concentrated on target detection and recognition. Each set of observations included both S-1 and S-2 results, i.e., the subject was required to respond from a position outside the simulated rain but viewing through it and then from within the simulated rain. The following variables and criterion measures were used:

Variable or Measure	S-Test
Independent variable	
Rainfall intensity	1, 2
Time (day or night)	1, 2
Glare versus no glare	1, 2
Ambient light	1, 2
Windshield wiper rate	2
Controlled variables	
Interior environment and fogging	1, 2
Target position (separation distance)	1, 2
Subject-vehicle position (separation distance)	1, 2
Glare-vehicle position	1, 2
Target presentation	1, 2
Original equipment windshield wipers and variable speed modification	2
Criterion measures	
Target detection	1, 2
Target recognition	1, 2
Visual acuity	1, 2
Legibility	1, 2

The rain intensity was adjusted and checked before the beginning of each test period. The glare vehicle was positioned (at night) with the visual acuity and legibility signs exactly 45.7 m (150 ft) from the subject vehicle. The artificial rain was begun, and the subject maneuvered the vehicle to a predesignated position (S-1) outside the rainfall. Instructions concerning the desired observations were read to the subject, whose view was then restricted by placing a cardboard shield across the front of the vehicle. The Landolt rings were repositioned, and the destination sign was changed. The shield was removed, and the subject was asked to respond. The test administrator recorded the experiment number, the time, the ambient light (during the day), the glare condition (at night), the rain intensity, the visual acuity, and the legibility (correct or incorrect).

The subject then maneuvered the vehicle to a new position (S-2) within the rainfall. The shield was replaced and a wiper rate was established and recorded. The Landolt rings and the destination sign were reset and the shield was removed. The subject was asked to respond in the same manner as before, and the same data were recorded. This procedure was repeated throughout the range of wiper rates and for both glare conditions. (The Landolt ring orientations, the destination-sign names, and the variations in wiper rates and glare conditions were all presented in random order.) The procedures for administering the tests involving target detection and recognition were essentially the same.

RESULTS

The analysis of the data began with a survey of the simple statistics; those results that appeared promising were then studied by an analysis of variance (ANOVA) and a regression analysis to develop dependencies.

S-1 Tests

The S-1 tests included visual acuity, probability of detection, and legibility measured at various rain rates [2.5, 5.1, 7.6, and 10.6 cm/h (1, 2, 3, and 4 in/h)] under both daytime and nighttime conditions. The simple statistics for visual acuity, measured in minutes of visual angle in daytime and dark conditions respectively, are given below (1 cm = 0.4 in).

	Rain Rate (cm/h)				
Daytime Statistic	0	2.5	5.1	7.6	10.2
Mean	1.058	1.563	1.400	1.972	2.688
Standard deviation	0.243	0.512	0.507	0.580	0.814
Low	1.000	1.000	1.000	1.000	1.000
High	2.000	2.000	2.000	2.500	5.000

	Rain Rate (cm/h)				
Nighttime Statistic	0	2.5	5.1	7.6	10.2
Mean	1.716	2.324	2.367	3.350	7.775
Standard deviation	0.479	0.868	0.482	1.587	2.775
Low	1.000	1.000	1.000	2.000	3.500
High	2.500	4.500	3.500	10.000	10.000

The visual acuity shows a definite increasing trend as the rain rate increases. The average values are plotted as a function of the rain rate in Figure 3. At night there was a large decrease (increase in the minimum visual angle) in visual acuity at higher rain rates although in daylight very high rain rates produced only a minor degradation of visual acuity. Quite obviously, rain does not significantly affect visual acuity if there is no windshield and water interface.

There is a seeming anomaly in the visual acuity at the 2.5-cm/h (1-in/h) rate. The simulator produced a greater proportion of drops smaller than 0.5 mm in diameter at the 2.5-cm/h (1-in/h) rate than at the 5.1-cm/h (2-in/h) rate, which is evidence that the drop size, rather than the rain rate, is the primary factor in reducing visibility (fog being the limiting case). Unfortunately, the net effect is not quantifiable with the data collected, as the variability is so high that any statistically significant effect is masked.

There was no attempt to perform a regression analysis on the S-1 data for visual acuity since the standard deviations showed a variability that increases nonlinearly over the range of rain rates. Also, the drop size must have a marked effect since the values of the standard deviation at the 2.5-cm/h (1-in/h) rate have no relation to the values at the 0 and 5.1-cm (2-in) rates. Consequently, the error variance of the random variable (the visual acuity) violates the requirement of uniformity in a manner that cannot be corrected by weighting, which obviates the validity of a linear regression analysis.

The S-1 day results, however, showed much less variability, and the results of the ANOVA of the visual acuity with the rain rate are shown in Table 1. The F-ratio is clearly significant. The Duncan multiple range test (Table 2) shows that the visual acuity is significantly different from zero for the 7.6 and 10.2-cm/h (3 and 4-in/h) rates and for the 5.1 and 10.2-cm/h (2 and 4-cm/h) rates. The visual acuity in up to 7.6-cm/h (3-in/h) rain rates is not significantly different from that in the clear condition. (The statistic used in the ANOVA tests the hypothesis of equal treatment means, i.e., H_0: $M_1 = M_2 = M_3 = M_4$, by testing the hypothesis of equal variance, i.e., H_0: $\sigma_1^2 \quad \sigma_2^2 \quad \sigma_3^2 \quad \sigma_4^2$. Therefore, the ANOVA results may mean only that the variances were different at the different rain rates. The Duncan multiple range test then relies on the error variance, which is considered uniform from treatment level to treatment level.

The detection tasks showed no perceptible degradation during the daytime rain condition. The nighttime condition, however, showed some interesting results. There was no particular pattern of detection or identification probability with the rain rate. Here, several variables that were not experimentally controlled affect the system. Specifically, although headlight glare would be expected to reduce target identification and detection, it actually increased the probability of detection at the 7.6-cm/h (3-in/h) rate. Not only does rain increase the specular reflection, and hence the disability glare, but the water in the atmosphere also causes increased backscatter, which possibly illuminates the object to be detected (3). Rain size and rate have a confounding effect on visibility that is dependent on the task to be performed.

S-2 Tests

The S-2 tests were conducted in the same manner as the S-1 tests except that the vehicle was in the rain. The most obvious result is the significant effect of water on the windshield. The probability of detecting the sign dropped significantly, and the probability of reading the sign dropped to zero at the lowest level of rain rate (Figures 4 and 5). These figures also show the effect of the windshield wiper; at even the lowest wiper speed the probability of detecting and reading the sign increases.

Under the nighttime condition, the detection of the sign is the same under both the glare and the no-glare conditions. However, the probability of reading the sign behaves anomalously. Under both the glare and the no-glare conditions, the probability of reading at the 2.5-cm/h (1-in/h) rate is less than that for the 5.1 and 7.6-cm/h (2 and 3-in/h) rates. The most reasonable explanation for this phenomenon is that of the drop-size distribution. The improved probability of reading the sign in the glare condition is the result of the illumination of the sign by the backscattered light.

The probability of detecting the individual targets and properly identifying them was essentially the same for all of the rain rates with the use of windshield wipers during the daytime condition. The nighttime condition, however, presented a different picture. The data for the probability of detection and proper identification showed a precipitous degradation between the 7.6 and 10.2-cm/h (3 and 4-in/h) rain rates. The curves in Figures 4 and 5, for the probability of detection and the probability of identification respectively, show the marked improvement of visibility given by windshield wipers. Between the 0 and 2.5-cm/h (1-in/h) rate the probability of detection dropped from almost 1 to less than 0.50. (There are no data points in this range to statistically support a hypothesis regarding the shape of the curve, but the function is probably a negative exponential.)

Thus, glare causes decreased detectability at night and reduces identification even more, although the differences are not significant in these data. The data at the 2.5-cm/h (1-in/h) rate appear to be an artifact but may be explained by the effect of the previous drop size. Apparently, the smaller drop-size distribution causes greater backscatter from both the approaching object and the vehicle, and this backscatter, while illuminating the sign, also causes an increased background illumination that reduces the contrast between the sign and the background.

The simple statistics for the visual acuity data for the S-2 daytime and nighttime simulator studies were calculated and are graphically shown in Figure 6. The improvement in visual acuity with the use of a wiper definitely indicates that water on the windshield is the most significant aspect of rain-reduced visibility. However, changes in the wiper speed have virtually no effect on visual acuity [wiper speeds below 25 cycles/min (cpm) were not investigated].

The design of the S-2 experiment lends itself to a three-way classification ANOVA, which is shown in Table 3. In this analysis, three levels of time (daytime, nighttime, and nighttime with glare), four levels of wiper speed (0, 25, 50, and 75 cpm), and five levels of rain rate [0, 2.5, 5.1, 7.6, and 10.2 cm/h (0, 1, 2, 3, and 4 in/h)] were used. Since no inferences about the population of rates were made, and since the other classification variables were discrete, a fixed-effect model was chosen for the analysis. All of the effects, including the interactions, were significant. The F-ratios were very significant, indicating that visual acuity is degraded

Figure 1. Rainfall simulator.

Figure 2. Landolt ring and legibility sign.

Figure 3. Visual acuity versus rain rate (S-1).

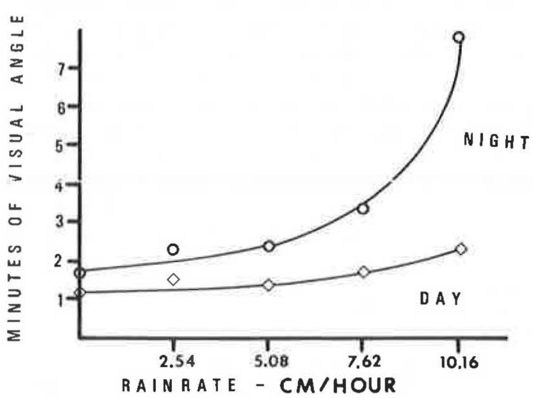


Table 1. Statistical analysis (S-1 day data).

Source	Degrees of Freedom	Sum of Squares	Mean Square	F-Ratio
Rate	4	7.233	1.808	12.056[a]
Error	10	1.500	0.150	–
Total	14	8.733	0.623	–

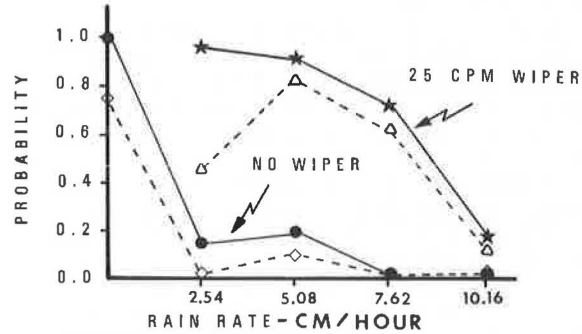
[a]Significant at the 0.01 level.

Table 2. Duncan multiple range test (S-1 day data).

Rain Rate (cm/h)	Mean of Visual Angle	Rain Rate (cm/h) 0	2.5	5.1	7.6	10.2
0	1.0	0	0.33	0.66	1.50[a]	1.83[a]
5.1	1.33	–	0	0.33	1.17	1.50
2.5	1.66	–	–	0	0.94	1.17
7.6	2.50	–	–	–	0	0.33
10.2	2.83	–	–	–	–	0

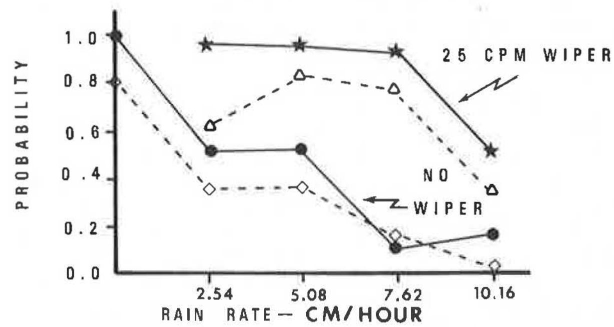
Note: 1 cm = 0.4 in.
[a]Significant at the 0.05 level.

Figure 4. Probability of detecting the sign (S-2 night).



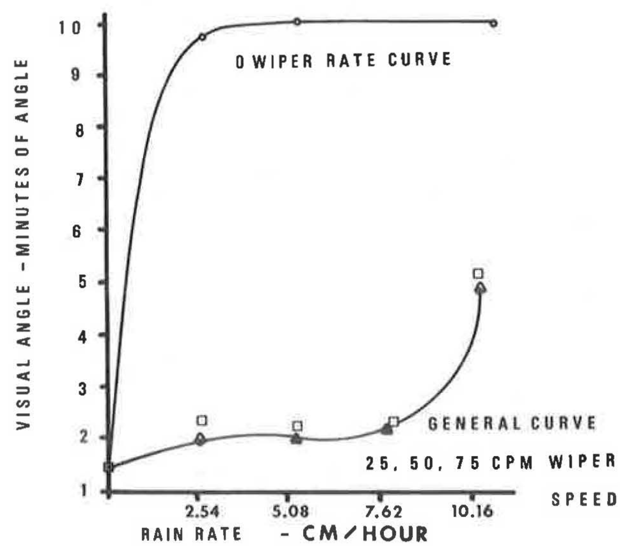
Note: Glare condition shown by dashed line.

Figure 5. Probability of reading the sign (S-2 night).



Note: Glare condition shown by dashed line.

Figure 6. Visual acuity versus rain rate (S-2).



by the time of day, the rain rate, and the wiper speed. The results of a Duncan multiple range test of these data are shown in Table 4. This test shows that the visual acuity at the 0-cpm wiper speed is significantly less than that at the 25, 50, and 75-cpm wiper speeds for all of the rain rates investigated. The differences in visual

Table 3. Analysis of variance (S-2 data).

Source	Degrees of Freedom	Sum of Squares	Mean Square	F-Ratio
Rate	4	1 609.66	402.41	204.27[a]
Wiper speed	3	4 673.87	1557.96	790.84[a]
Time	2	273.38	136.69	69.38[a]
Rate[a], wiper speed	9	2 445.52	271.72	137.93[a]
Rate[a], time	8	432.96	54.12	27.47[a]
Wiper speed[a], time	6	198.44	33.07	16.79[a]
Rate[a], time[a], wiper speed	18	143.36	7.96	4.04[b]
Error	871	1 718.02	1.97	
Total	921	11 495.22	12.48	

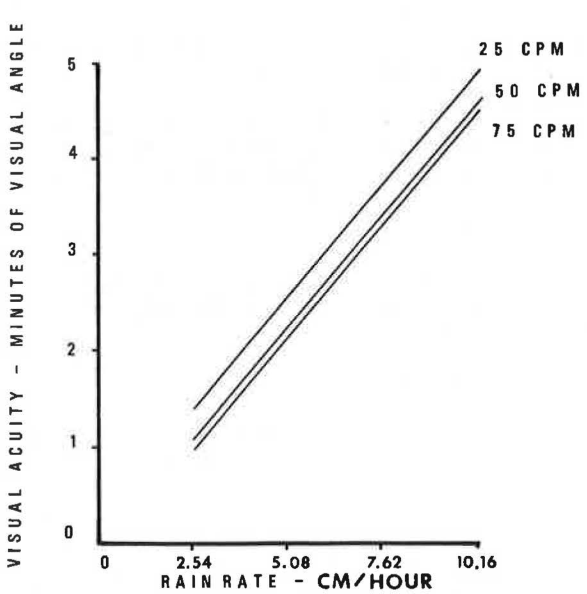
[a]Significant at the 0.0001 level. [b]Significant at the 0.01 level.

Table 4. Duncan multiple range test (S-2 data).

Visual Acuity	V_0	V_{25}	V_{50}	V_{75}
2.5-cm/h rain rate				
V_0 = 9.76	0	7.39[a]	7.76[a]	7.77[a]
V_{25} = 2.37	—	0	0.37	—
V_{50} = 2.00	—	—	0	—
V_{75} = 1.99	—	—	—	0
5.1-cm/h rain rate				
V_0 = 7.98	0	5.71[a]	5.85[a]	7.73[a]
V_{25} = 2.27	—	0	0.14	0.15[b]
V_{50} = 2.13	—	—	0	0.01
V_{75} = 2.12	—	—	—	0
7.6-cm/h rain rate				
V_0 = 9.83	0	5.71[a]	5.85[a]	5.86[a]
V_{25} = 2.40	—	0	0.14	0.15[b]
V_{50} = 2.12	—	—	0	0.01
V_{75} = 2.10	—	—	—	0
10.2-cm/h rain rate				
V_0 = 10.00	0	5.71[a]	5.85[a]	5.86[a]
V_{25} = 5.63	—	0	0.14	0.15[b]
V_{50} = 5.28	—	—	0	0.01
V_{75} = 5.76	—	—	—	0

Notes: 1 cm/h = 0.4 in/h.
V = visual acuity in minutes of visual angle.
[a]Significant at 0.01 level. [b]Significant at 0.05 level.

Figure 7. Prediction of visual acuity at various wiper speeds by the regression equation.



acuity among the 25, 50, and 75-cpm wiper speeds were either not significant or were barely significant at the 0.05 level. Thus, wiper speeds greater than 25 cpm do not significantly improve visual acuity at rain rates of up to 10.2 cm/h (4 in/h).

After the data were altered to give visual degradation by improving the absolute threshold angle in clear air, further analysis of the visual acuity data for the S-2 daytime tests by multiple linear regression techniques gave Equation 1.

$$VA = 0.415\,25r + 0.755\,59W_{25} + 0.597\,05W_{50} + 0.584\,86W_{75} \qquad (1)$$

where

VA = degradation of threshold visual angle in minutes,

Figure 8. Cross section of vehicle windshield at visual centerline.

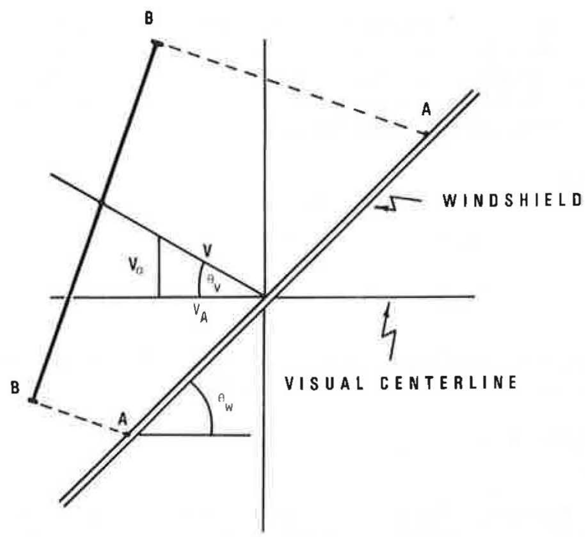
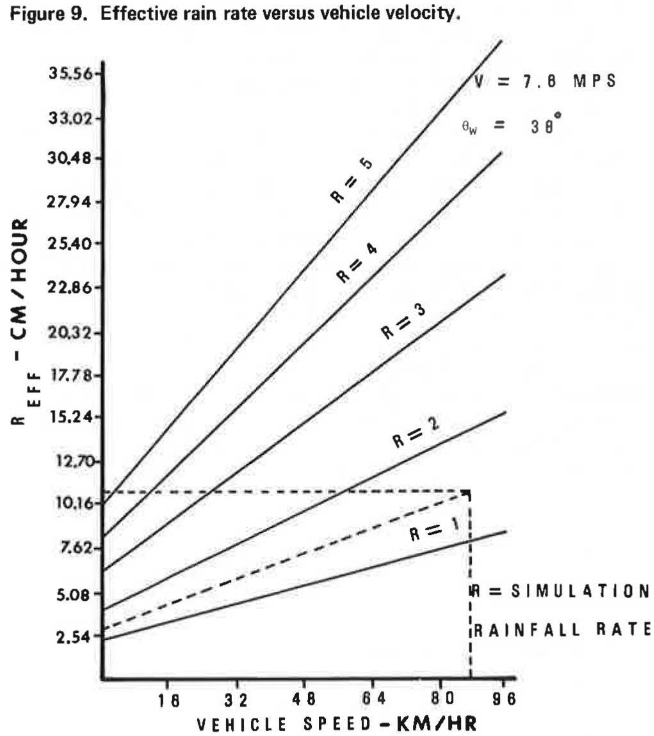


Figure 9. Effective rain rate versus vehicle velocity.



r = instantaneous rain rate, and
W_{25}, W_{50}, W_{75} = one when wiper speed is 25, 50, or 75 cpm respectively and zero otherwise.

The model uses three dummy variables for the three conditions of wiper speed (W_{25}, W_{50}, and W_{75} equal one when the wiper speed is 25, 50, and 75 cpm respectively; otherwise zero). The regression analysis of variance (shown below) shows that the model is highly significant, with a coefficient of multiple regression of 0.914, and that all of the parameters are also highly significant (α = 0.0001).

Source	Degrees of Freedom	Sum of Squares	Mean Square	F-Ratio	Prob>F
Regression			94.41	315.96	0.0001
Rain	1	368.54	368.54	1233.40	0.0001
Wiper rate	3	9.09	3.03	10.14	0.0001
Error	119	35.56	0.299	–	–
Total	123	413.19	–	–	–

The test for the significance of the regression coefficient for wiper speed (shown below) shows that the 75-cpm rate is significantly different from the 50 and 75-cpm rates, which leads to the conclusion that wiper speeds need not be greater than 50 cpm.

Wiper Speed (cpm)	t for $H_0B = 0$	Prob>T
0	9.33	0.0001
25	5.37	0.0001
50	4.24	0.0001
75	4.16	0.0001

Figure 7 shows the prediction of visual acuity at various wiper speeds by the regression equation.

Relation Between Simulated and Natural Rainfall

The vital link relating the simulator studies and natural rainfall is the effective rain rate. That is, the simulator studies involve static tests in which the rain falls directly on the windshield, but in actual driving conditions, the rainfall on the windshield is a function of the vehicle velocity: Simulator rain rates correspond to much lower effective rain rates encountered in the dynamic rain environment.

There is a direct relation between the effective rate of rainfall on the windshield and the vehicle speed, the actual rainfall rate, and the rake angle of the windshield. The effective rate of rainfall on the windshield can be estimated by summing the rates due to falling raindrops and to the forward (horizontal) motion of the vehicle (4). If the effects of the aerodynamics of vehicle design and the raindrops splattering on the vehicle are neglected, and if it is assumed that all raindrops in a given rain are falling straight downward at the same velocity, then the effective rain rate (r_{eff}) is defined as the static rain intensity necessary to produce the same amount of water on the windshield as would be encountered in an actual rain of intensity (r) in a vehicle traveling at velocity (v):

$$r_{eff} = f(r, v_a, v_0, \theta_w) \tag{2}$$

where

r = actual rain rate,
v_a = vehicle velocity,
v_o = terminal velocity of raindrops, and
θ_w = windshield rake angle.

The rain rate is defined as the depth of water falling on a unit area in a given time interval (typically an hour). If the unit area is considered as moving through a stationary, water-filled atmosphere, the relation between r and r_{eff} can be derived. Figure 8 shows the resultant velocity vector for the windshield moving at velocity (v_a) during a rainfall with a terminal velocity (v_o). The magnitude of this vector is given by

$$v = (v_0^2 + v_a^2)^{1/2} \tag{3}$$

and the angle is given by

$$\theta_v = \tan^{-1}(v_0/v_a) \tag{4}$$

Consider the plane unit area ($\overline{BB}$) normal to the resultant velocity vector ($\vec{v}$). The effective rain rate, if that plane moves with velocity (v), is

$$r'_{eff} = (v/v_0)\,r \tag{5}$$

The plane ($\overline{BB}$) projects onto the windshield to form the plane ($\overline{AA}$) Thus, the effective rain rate is r'_{eff} reduced by the ratio of the unit area at $\overline{BB}$ divided by its projection on the windshield. This can be reduced to the ratio of $\overline{BB}$ to $\overline{AA}$ or

$$\lambda = \overline{BB}/\overline{AA} = \cos[90 - (\theta_v + \theta_w)] \tag{6}$$

Since $\overline{BB}$ is a unit area, this reduces to

$$\lambda = \sin\theta_w \cos\theta_v + \cos\theta_w \sin\theta_v \tag{7}$$

or (from Equation 4)

$$\lambda = \sin\theta_w \cos[\tan^{-1}(v_0/v_a)] + \cos\theta_w \sin[\tan^{-1}(v_0/v_a)] \tag{8}$$

The effective rain rate is now given by the following relation:

$$r_{eff} = \lambda r'_{err} = \lambda[(v_0^2 + v_a^2)^{1/2}/v_0]\,r \tag{9}$$

This becomes (from Equation 8)

$$r_{eff} = \{[v_a/(v_0^2 + v_a^2)^{1/2}]\sin\theta_w + [v_0/(v_0^2 + v_a^2)^{1/2}]\cos\theta_w\}\ (v_0^2 + v_a^2)^{1/2}/v_0 \tag{10}$$

which reduces to

$$r_{eff} = r[(v_a/v_0)\sin\theta_w + \cos\theta_w] \tag{11}$$

The veracity of this relation was checked by investigating the following limiting cases:

1. A windshield with a rake angle of 0 deg at zero velocity (v_a = 0) should have an effective rain rate equal to the actual rain rate. The relation for r_{eff} shows that

$$r_{eff} = r[(0/v_0)\sin 0^\circ + \cos 0^\circ] = r \tag{12}$$

2. A windshield with a rake angle of 90 deg at zero velocity (v_a = 0) should have an effective rain rate of zero (r = 0). The relation for r_{eff} shows that

$$r_{eff} = r[(0/v_0)\sin 90^\circ + \cos 90^\circ] = 0 \tag{13}$$

3. A windshield with a rake angle of 90 deg at velocity v_a should have an effective rain rate of

$$r_{eff} = r[(v_a/v_0)\sin 90^\circ + \cos 90^\circ] = r(v_a/v_0) \tag{14}$$

Thus, as the vehicle velocity increases, the effective rain rate increases. Further, as the terminal velocity decreases, the rain rate increases because the amount of water in the air at any instantaneous time also increases. And at 0-deg windshield rake angle an increase in velocity has no effect on effective rain rate.

Figure 9 shows a plot of effective rain rates versus vehicle speed for selected rainfall rates. This plot makes two significant assumptions: (a) that the vehicle velocity vector and the rainfall were at 90 deg to each other and (b) that the effects of wind could be ignored. The curves show that the rain produced by the simulator accurately reflects rain rates that are typically encountered. For example, to simulate the condition of a vehicle having a velocity of 88 km/h (55 mph) in a rainfall of 3.8 cm/h (1.5 in/h) requires a static rain rate of 10.80 cm/h (4.25 in/h). The dotted lines in the figure show this relation.

CONCLUSIONS

The significant results of this research can be summarized as follows:

1. During rain conditions, the primary factor that reduces visibility is the film of water on the windshield, which impairs vision by reducing the optical resolution. The S-1 studies, when compared to the S-2 studies (no rain on the windshield versus rain on the windshield), demonstrate this point. At a 2.5-cm/h (1-in/h) simulator rain rate [equivalent to a 0.75-cm/h (0.30-in/h) effective rate at 88 km/h (55 mph)], vision through the windshield is reduced to the point that acuity decreases to 10 min of visual arc, which corresponds to a static visual acuity of 20/200. However, the daylight visual acuity through a 10.2-cm/h (4-in/h) simulator rain, with no water on the windshield, produced a visual degradation equivalent to only 2.5 min of visual arc, which corresponds to a static visual acuity of 20/50.
2. The simulator results showed a precipitous decrease in the detection and identification of pertinent targets (i.e., a man or an automobile) between the 5.1 and 10.2-cm (2 and 4-in/h) simulated rain rates.
3. Windshield wipers restore visual acuity to approximately the same level as would be expected if the vehicle remained outside the rain and the driver looked through it. Higher windshield-wiper speeds do not significantly improve visibility at speeds above 50 cpm.
4. A regression model of visual degradation in terms of the increase in threshold visual angle as a function of the rain rate is given by Equation 1.
5. There are significant interactions between rain and the glare from oncoming vehicles.
6. Raindrop size distribution is a significant factor in visibility reduction, especially at low levels of illumination. A concentration of smaller drop sizes, i.e., those less than 0.5 mm in diameter, causes serious visual degradation through reduction of contrast and the decrease in the quality of the texture background.

ACKNOWLEDGMENT

The research reported in this paper was sponsored by the National Highway Traffic Safety Administration. The contents reflect the views of the authors only, who are responsible for the facts and the accuracy of the data presented.

REFERENCES

1. D. L. Ivey, E. K. Lehitpuu, and J. W. Button. Rainfall and Visibility—The View From Behind the Wheel. Texas Transportation Institute, Texas A&M Univ., College Station, Rept. 135-3, 1975.
2. W. E. K. Middleton. Vision Through the Atmosphere. Univ. of Toronto Press, 1952.
3. A. Walvolgel. The N_o-Jump of Raindrops Spectra. Atmospharenphysik ETH, Zurich, 1974.
4. F. A. Berry. Handbook of Meteorology. McGraw-Hill, New York, 1945.

Publication of this paper sponsored by Committee on Visibility.

Computer Program for Roadway Lighting

F. W. Jung and C. Blamey, Research and Development Division, Ontario Ministry of Transportation and Communications

The development of a computer program for the design and evaluation of fixed highway lighting is reported. The program calculates the illuminance, luminance, and disability veiling brightness in each lane at specified grid points on the road surface for regular, straight rows of luminaires, for a straight highway up to six lanes wide. Isoilluminance and isoluminance diagrams can also be obtained. The program can be used as a design tool in the following way: For a chosen road geometry and a selected luminaire type, the designer can determine the performance of a proposed lighting design by calculating the relevant performance measures and comparing the results with the current accepted, or the proposed new standards. Many different designs can be rigorously evaluated in a short time. In conjunction with photometric measurements, the program was used to evaluate the performance of the existing design on the Toronto Bypass. Lighting designs based on calculations of luminance and disability veiling brightness are preferable to those based on illuminance because nighttime visibility is determined by the former rather than the latter.

Modern electronic computer methods are entering the field of outdoor lighting and assisting and improving the design and management of lighting systems. This paper presents a model for a computer program that combines the design tasks of luminaire selection, performance evaluation, and, at a later stage, economic comparison of various alternative systems. The domain of this model is limited to straight, regular systems of roadway lighting, but similar models can be used for other lighting systems, such as parking lots, shopping plazas, or curves and intersections of highways (although the higher costs of developing these programs may be justified only if their potential users join in the effort).

Lighting design by computer methods is cost-effective for two reasons: First, there is a saving in labor costs when computers are used efficiently and, second, more efficient designs will be developed because computer methods permit more rigorous analysis of the performance of more alternatives than is possible with conventional methods. Thus, future lighting systems may have improved luminance uniformity on the street or roadway and also reduced consumption of electrical energy.

There is another long-range benefit from a computational approach. The computer program described here has been modeled with performance parameters that are oriented toward the visual task of night driving. The use of these parameters can avoid overdesign of lighting systems if standards are adopted that are more relevant to the night-driving task than those traditionally used. For example, lighting systems can be designed for contrast sensitivity by using background luminance rather than roadway illuminance. Or, the system might be designed for an acceptable glare level rather than by using cutoff specifications. The use of the computer program will therefore permit lower illuminance levels that have less glare and are more uniform.

In developing this approach to roadway-lighting design, previous results in the fields of visibility, pavement reflectance, and glare have been considered. Much of this research has been done in Europe, where the problems of energy conservation and the quality of roadway lighting have been more acute than in North America.

SYSTEM LAYOUT AND INPUT SUBPROGRAM

The overall layout of the illumination-design program is shown in Figure 1. The first part of the program (numbers 1 through 7) contains the technical evaluation of alternative lighting designs and is the subject of this paper. The second part of the program contains an economic cost model and will be added later. The design procedure is as follows: A list of suitable luminaires and design arrangements for a particular project (i.e., for a cross section of a given road) is assembled by means of a subprogram (number 2 of Figure 1). The designer establishes the input data for the road section and then, sitting at a terminal, selects various luminaires and suitable arrangements by typing values and code numbers in response to questions asked by the computer program. The illumination levels are determined, and the spacings or uniformities are calculated by the computer as in a conventional design method, except that the computer uses digitized photometric data stored in a luminaire data bank. Whenever a suitable luminaire is selected and a design that has a sufficient average level and uniformity of illuminance is found, these data are added to the input for the part of the program designated Illum 1, which calculates the performance parameters and evaluates the performance of each design.

In this first subprogram, the uniformity is calculated as the ratio of the average to the minimum illuminance. The minimum value of illuminance is chosen from a limited number of point-by-point calculations that use digitized luminous-intensity data for each type of luminaire. The average level of the illuminance or the spacing is computed on the basis of digitized data for the coefficient of utilization. Thus, computerized forms of traditional design procedures (numbers 1 and 2) are used to preselect feasible luminaires and arrangements for the more rigorous performance-evaluation subprogram Illum 1 (number 4 in Figure 1).

The input subprogram and the Illum 1 subprogram use the same data bank input for photometric luminaire data. The most important data needed are the light distribution of the luminaires, i.e., the luminous-intensity distribution function [$I(\gamma, \phi)$], which is usually available on photometric data sheets issued by the manufacturers. Figure 2 illustrates the variable angles (ϕ and γ) of the luminaire-intensity function. These angles are defined by the equations below.

$$\phi = \arc\tan\,[a/(b - o)] \tag{1}$$

$$\gamma = \arc\tan\left\{[a^2 + (b - o)^2]^{1/2}/h\right\} \tag{2}$$

$$E_p = [I(\phi, \gamma)\cos^3\gamma]/h^2 \tag{3}$$

The manufacturers' data sheets usually contain the function (I) in the form of various diagrams, but the computerized method requires a format in which I is given in tabular form as a two-dimensional matrix corresponding to the two variables, the horizontal angle (ϕ) and the vertical angle (γ). (Between the discrete values given by the matrix or table, other values can be determined by parabolic interpolation.) A format for a symmetrical luminaire [$I(\gamma, \phi) = I(\gamma, -\phi)$] is given in Figure 3 for vertical angles below the horizon ($\gamma \leq 90$). The format for the coefficient of utilization is given in terms of the ratios (b - o):h for the street side or o:h for the house side, where either ratio can vary between 0.0 and 6.0.

ILLUM 1: CALCULATION OF PERFORMANCE PARAMETERS, ILLUMINANCE, AND DISABILITY VEILING BRIGHTNESS OR GLARE

Single values of illuminance, luminance, and disability veiling brightness or glare (DVB) are calculated for selected grid points on, or over, the road surface for one section between a repetitive arrangement of luminaires. The grid points represent the point (P) on the road surface in Figure 2, or the position of the driver's eyes as shown in Figure 4. The following equations are used for the calculation of the illuminance and the DVB (1).

$$\Theta = \arc\tan\,[(h - e)^2 + (b - o)^2]^{1/2}/d \tag{4}$$

$$-\phi = \arc\tan\,[d/(b - o)] \tag{5}$$

$$\gamma = \arc\tan\left\{[d^2 + (b - o)^2]^{1/2}/(h - e)\right\} \tag{6}$$

$$R = (h - e)/\cos\gamma \tag{7}$$

$$E_\nu = [I(\phi, \gamma)\cos\Theta]/R^2 \tag{8}$$

$$DVB = 10E_\nu/\Theta^2 \tag{9}$$

The arrays of single values are then added and averaged, or scanned for maximum or minimum values, as required.

LUMINANCE OR REFLECTED LIGHT

Luminance is calculated from the corresponding illuminance values for the same grid points (P), but only the portion of the light that is reflected toward the driver's eyes is considered. As shown in Figure 5, the illuminance contribution (E_p) from each luminaire is multiplied by a coefficient (q) that depends on the light-reflection properties of the pavement surface (4). For each driver position, or each lane, the calculated luminance arrays are different (unlike the illuminance arrays, which remain the same).

The reflection of light from a road surface ranges from complete specularity (the mirror effect), when the surface is flooded with water, to almost complete diffusion for a nonglossy, dry pavement. However, dry or almost dry conditions prevail most of the time, and the increase in glossiness of damp pavements usually in-

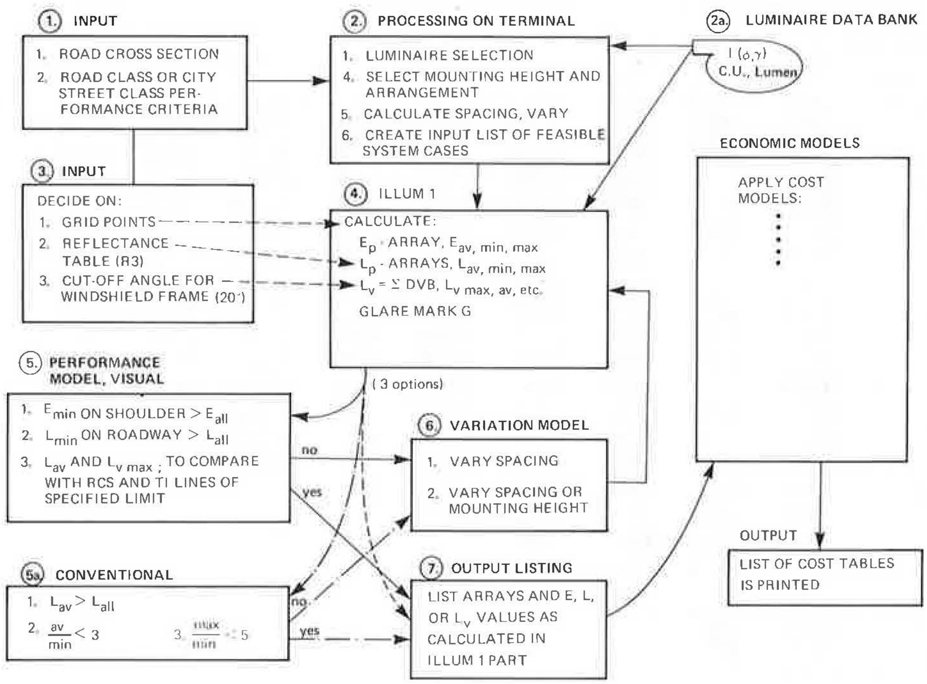


Figure 1. Overall flow diagram for illumination-design program.

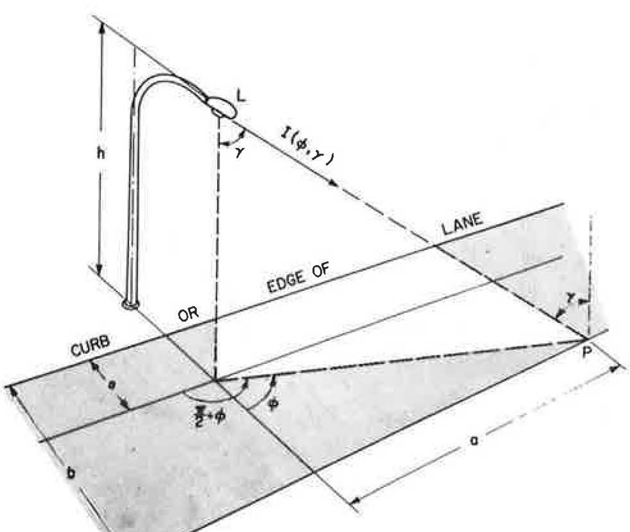


Figure 2. Illuminance.

creases both the average luminance (5) and the difference between the average and minimum luminances. Since it is difficult to include all of these factors in the calculations, highway lighting is usually designed and evaluated for dry (to be eventually supplemented by inclusion of moderately wet) pavements, which permits classification of the degree of glossiness into a few standard categories. The number of categories would increase considerably if moist pavements were included.

On wet pavements, visibility, although diminished, is available from the familiar blurred and streaky image of the reflected luminaires. Under these conditions, laterally extended light sources, such as fluorescent or low-pressure sodium-vapor luminaires, installed above the roadway improve visibility because they generate wider streaks of blurred images.

As illustrated in Figure 5, the luminance coefficient is a function of four angles (α, β, γ, and δ) and of the average luminance coefficient (q_0), which depends on the color of the pavement surface. Thus,

$$L = q_0 \times [\bar{q}(\alpha, \beta, \gamma, \delta)] \times E_p \tag{10}$$

where

q_0 = average luminance coefficient derived from a specified road area,
$\bar{q}$ = luminance-coefficient function for a tabulated q_0 that is a function of the angles α, β, γ, and δ, and,
E_p = illuminance.

The luminance coefficient (q) is defined as the factor by which the illumination (E_p) must be multiplied to obtain the luminance (L). The luminance values must be calculated for each luminaire and then summed. The luminance created by a luminaire at point i is

$$L = q_0 \times [\bar{q}(\alpha_i, \beta_i, \gamma_i, \delta_i)] \times [E_p(\phi_i, \gamma_i)] \tag{10a}$$

If the values for n luminaires are added, Equation 3 is substituted for E_p, the influence of δ is neglected, and $\alpha = 1°$, the following equation for the luminance can be derived.

$$L = \sum_{i=1}^{n} \left\{ q_0 \times \bar{q}(\beta_i, \gamma_i) \times [I(\phi_i, \gamma_i) \times \cos^3 \gamma_i]/h^2 \right\} \tag{11}$$

Standard reflectance tables (4) have been established in the form of reduced coefficients ($R = \bar{q} \cos^3 \gamma$) for $\alpha = 1°$, which simplifies the reflectance measurements. The combination $R = \bar{q} \cos^3 \gamma$ leads to table values of R that decrease with increasing γ or tan γ, whereas the pure reflectance function ($\bar{q}$) or (q) alone increases greatly [Figure 6 (6)].

The number of luminaires (n) to be included are those within a longitudinal distance of 12 h beyond P; additional luminaires beyond this range contribute insignificantly. By the substitution of $R = \bar{q} \times \cos^3 \gamma$, Equation 11 can be rewritten as

Figure 3. Format for luminous-intensity distribution function.

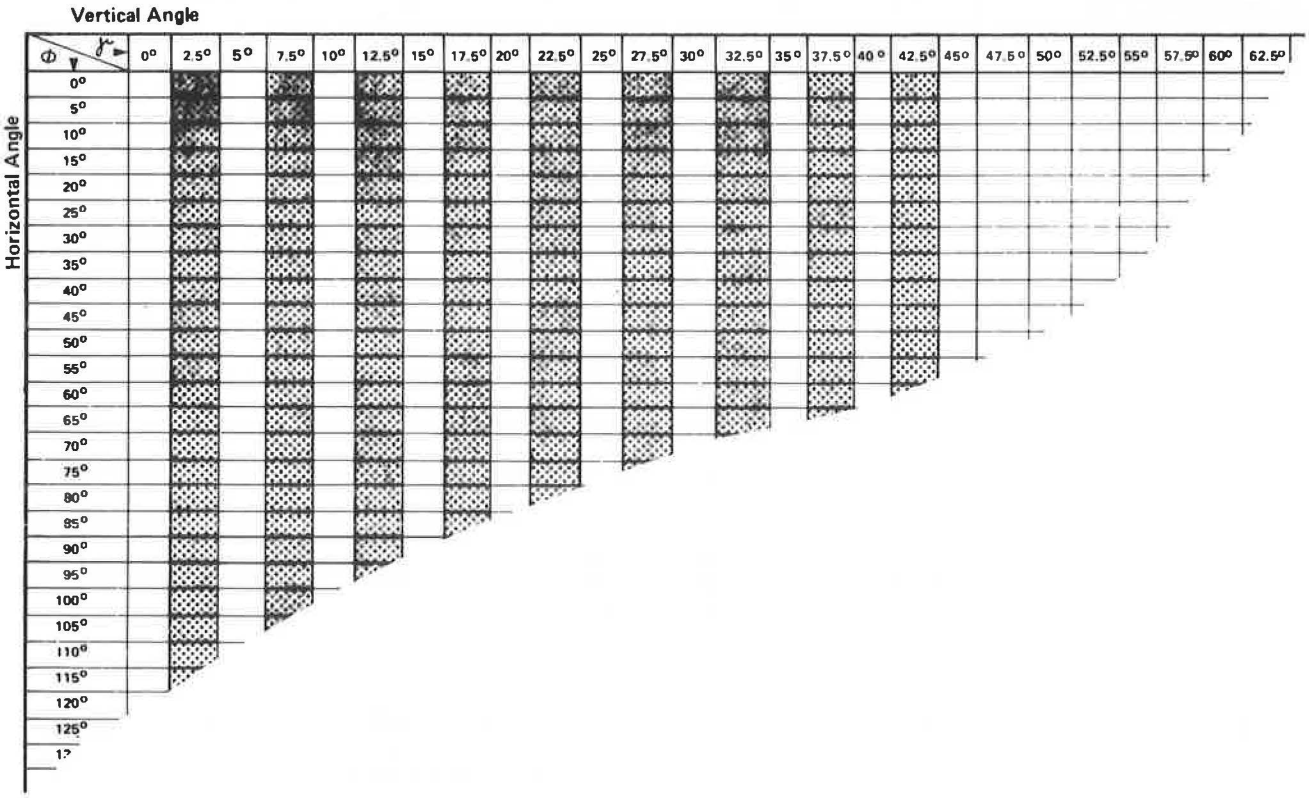
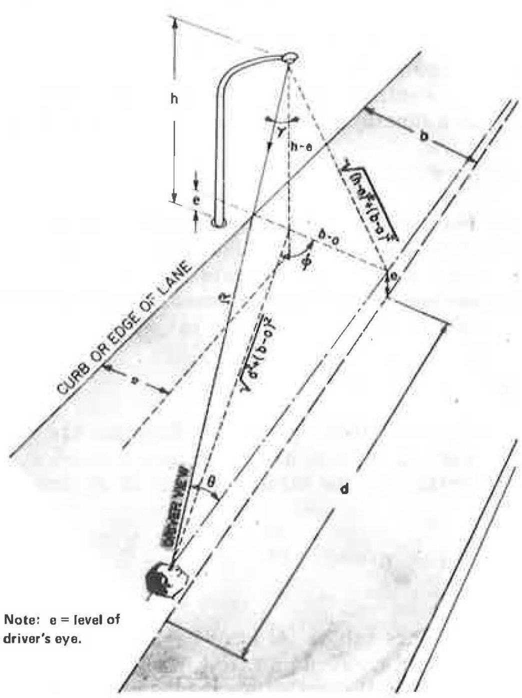
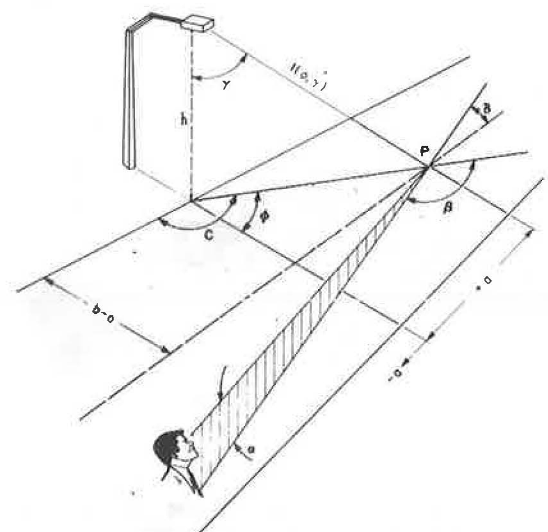


Figure 4. Disability veiling brightness.



Note: e = level of driver's eye.

Figure 5. Luminance.



$$L = \sum_{i=1}^{n} \left\{ [I(\phi_i, \gamma_i) \times R(\beta_i, \tan \gamma_i)] / h^2 \times q_0 \right\} \qquad (12)$$

An example of an abridged R-table, which tabulates values of R versus β and tan γ, is given in Table 1. Other examples are those of the Commission Internationale de l'eclairage (4) and Erbay (7).

Equations 11 and 12 represent a point-by-point method of calculating the luminance of a road surface as it appears to a driver in a particular lane who is looking ahead 90 m (300 ft). The first calculations must be carried out for all points on a perpendicular line across the pavement at this distance ahead of the driver. The next calculations assume that the driver has moved forward and is now looking at a line approximately 6.1 or 9.2 m (20 or 30 ft) ahead of the original line. Moving ahead in this way, the driver is assumed to maintain a constant, standard viewing angle of $\alpha = 1°$. Thus, all of the grid points on a road surface have as many arrays of luminance values as there are lanes for a driver to use, and the luminance values at these points are dynamic values of brightness successively reflected to the driver as he or she moves along. These values are not exactly the same as those seen by an observer from a stationary position.

OUTPUT OF PERFORMANCE VALUES

At this point, printouts for the calculated performance parameters—one array (per road side) of illuminance values for the specified grid points, arrays of luminance values for the grid points and each lane position, and one row of DVB values for each lane—can be obtained. Optionally, the array printouts can be converted into isoilluminance and isoluminance diagrams for more con-

Figure 6. Reflectance function.

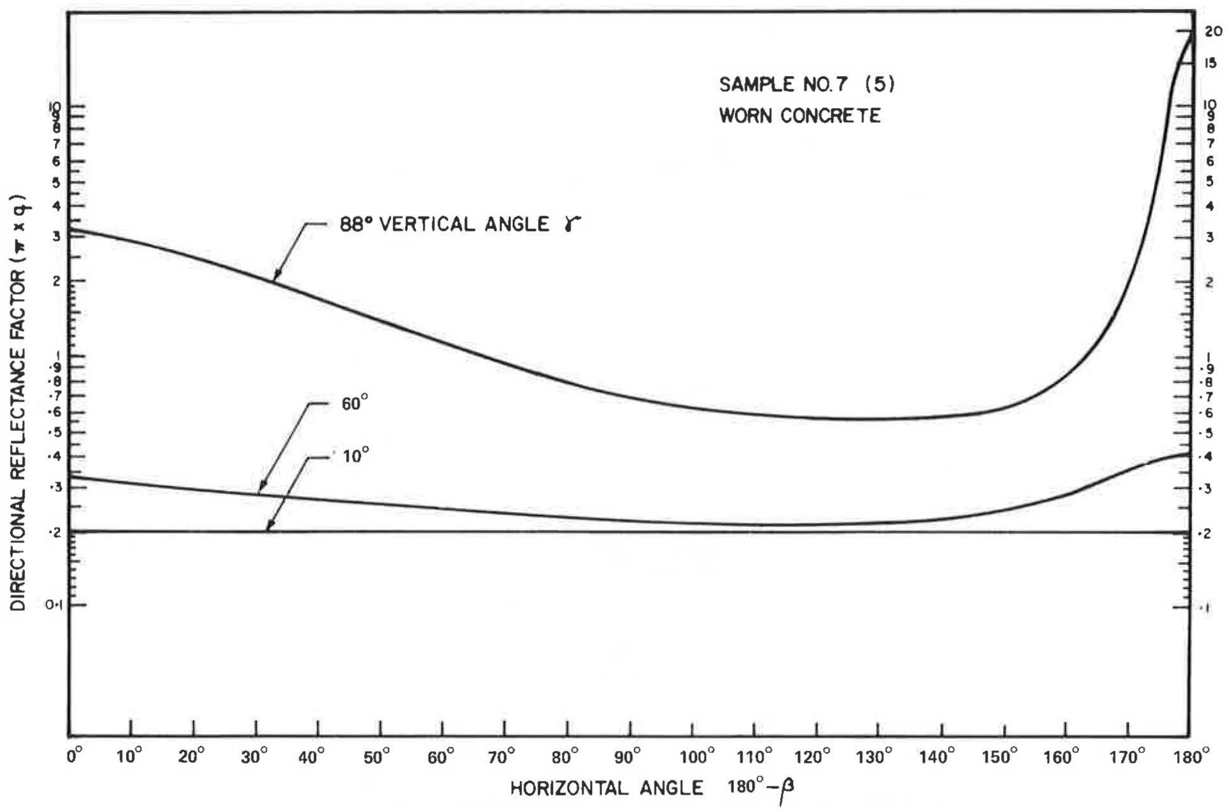


Table 1. Standard reflectance table R-3 (abridged).

	β[b]													
Tan γ[a]	0	2	5	10	15	25	35	45	60	75	90	120	150	180
0.00	426	426	426	426	426	426	426	426	426	426	426	426	426	426
0.50	498	498	491	491	471	445	419	380	367	321	295	288	288	282
1.00	524	524	511	472	400	328	262	203	197	157	144	144	144	144
1.50	511	504	472	386	314	210	144	118	109	89	83	86	86	89
2.00	472	465	406	275	197	119	89	69	62	54	48	51	52	55
2.50	419	406	321	183	124	77	55	43	39	34	31	34	35	37
3.00	367	341	236	123	76	45	33	26	24	21	20	22	24	25
3.50	314	282	177	86	51	31	24	18	16	14	13	16	17	18
4.00	275	236	131	62	38	24	17	13	12	0	0	12	13	14
4.50	236	197	106	45	29	17	13	10	0	0	0	9	10	12
5.00	210	157	85	34	24	13	10	9	0	0	0	0	0	0
5.50	183	131	68	26	20	10	8	0	0	0	0	0	0	0
6.00	164	111	52	21	16	9	7	0	0	0	0	0	0	0
6.50	151	98	43	16	12	8	5	0	0	0	0	0	0	0
7.00	138	86	35	12	9	7	0	0	0	0	0	0	0	0
7.50	128	76	30	10	8	5	0	0	0	0	0	0	0	0
8.00	121	68	25	9	7	4	0	0	0	0	0	0	0	0
8.50	113	60	21	8	5	4	0	0	0	0	0	0	0	0
9.00	106	55	17	7	5	3	0	0	0	0	0	0	0	0
9.50	100	50	14	5	4	3	0	0	0	0	0	0	0	0
10.00	94	46	13	5	4	3	0	0	0	0	0	0	0	0
10.50	89	42	12	4	3	0	0	0	0	0	0	0	0	0
11.00	85	38	10	4	3	0	0	0	0	0	0	0	0	0
11.50	81	34	9	4	3	0	0	0	0	0	0	0	0	0
12.00	77	31	8	3	3	0	0	0	0	0	0	0	0	0

[a] Tan γ = R/H values corresponding to the listed numbers of reflection values.
[b] All values of β have been multiplied by 1000.

Figure 7. Example graph plots.

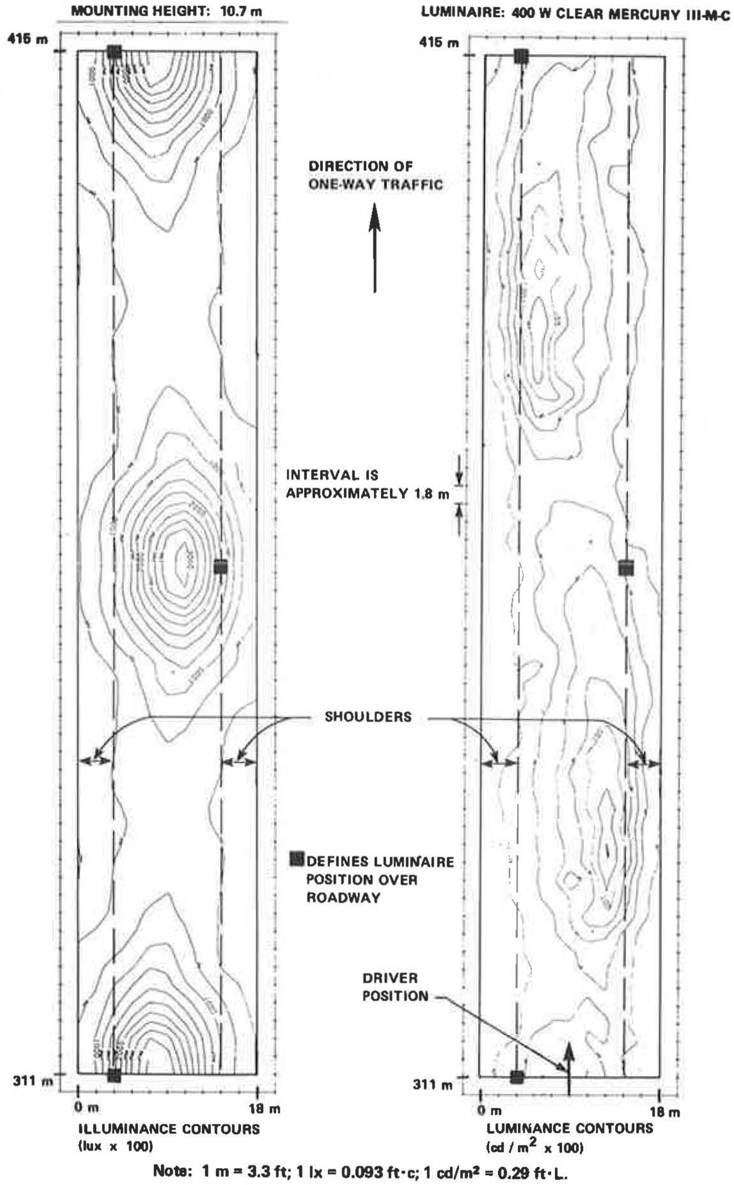


Figure 8. Diagram of visibility criteria.

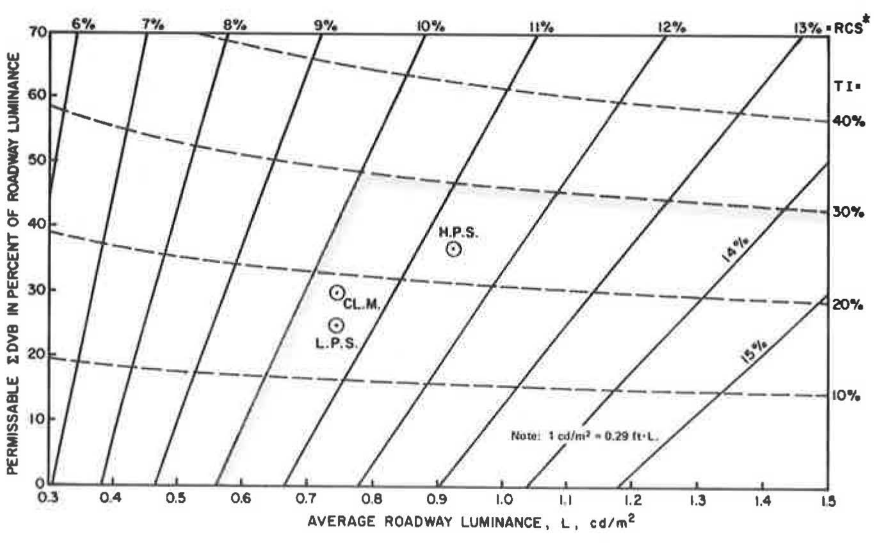
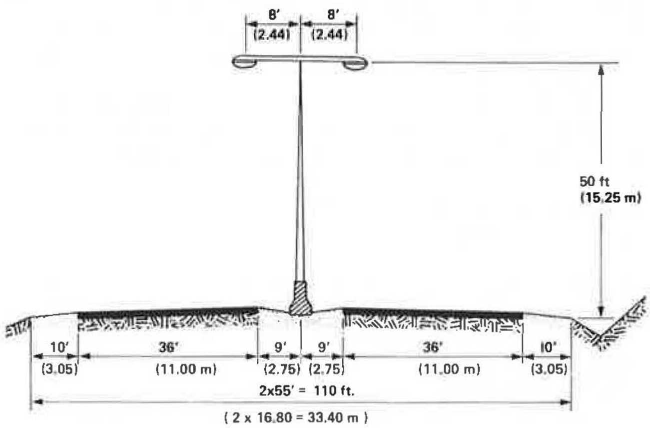


Figure 9. Design example.



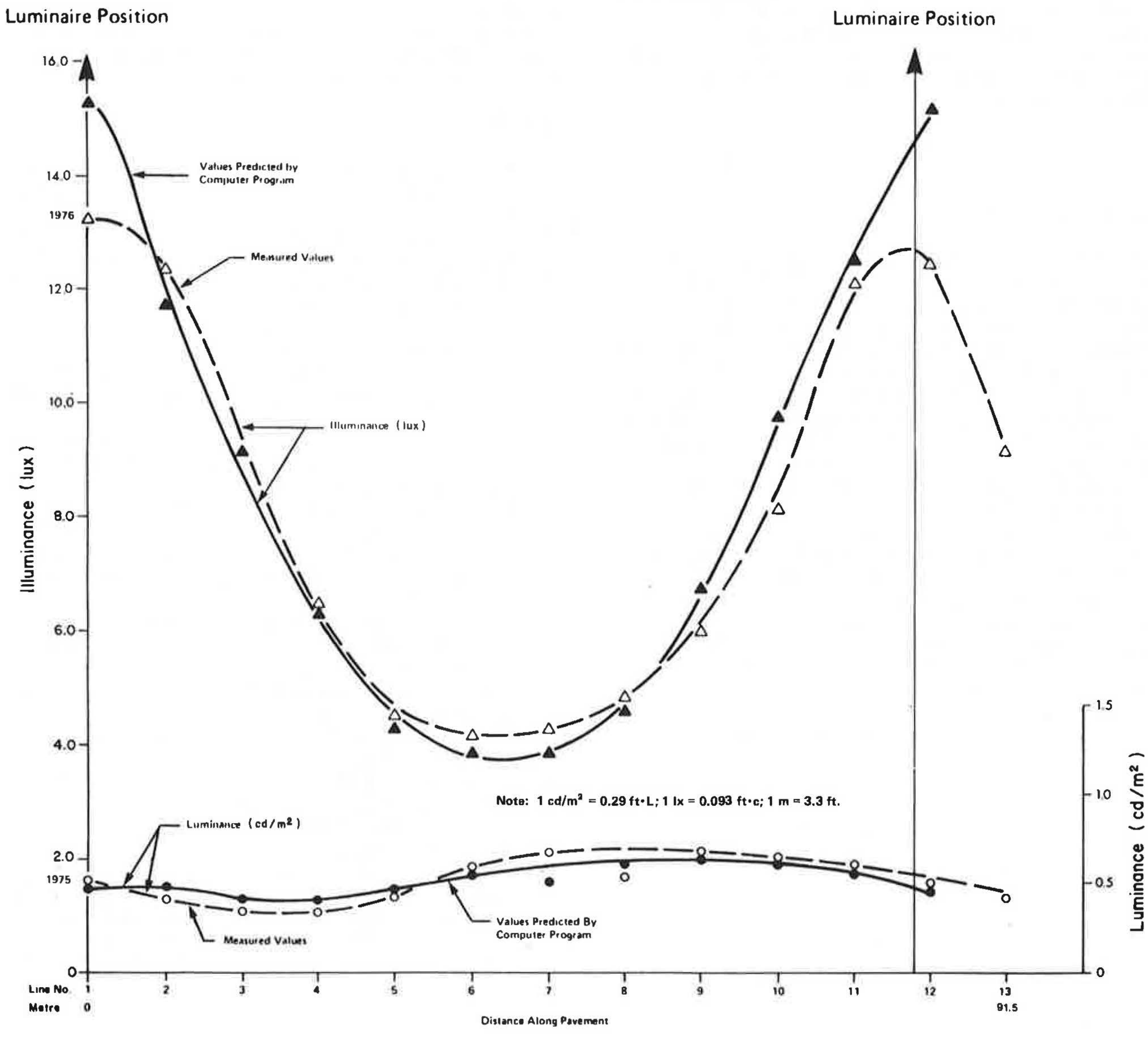
venient study by using a separate computer program. An example for a simple three-lane road is shown in Figure 7.

MODELING OF PERFORMANCE EVALUATION

The performance of a design traditionally has been evaluated by specifying limiting conditions for average values, uniformity ratios, or maximum and minimum values of the illuminance or incident light on the road surface, and correcting for glare by rigid cutoff specifications. This approach wastes an enormous amount of electrical energy, since much of the light is radiated toward places where it is not really needed. Traditional lighting standards are more concerned with maintaining good visibility than with saving energy.

The cost of lighting installations and the use of elec-

Figure 10. Comparison of measured and calculated values of illuminance and luminance (test area 4, bituminous overlay).



tric energy can both be reduced without reducing the level of service by redefining lighting standards in terms of the visual requirements of the night-driving task. The computer program, as it has been developed, can be modeled with traditional standards, or the performance evaluation could be modeled with the concept of relative contrast sensitivity (RCS), as discussed by Jung in the following paper. The RCS provided by fixed lighting at any important spot on the road surface must be larger than the specified minimum value required for the visual task of night driving. The following equation for RCS has been derived from standard values for lighting performance (8).

$$RCS = 13.7(L - 0.06)^{1/2} \quad (13)$$

This equation is valid for a luminance range of 0.15 to 2.5 cd/m^2 (0.044 to 0.73 ft•c) and for glare-free lighting installations. This value of RCS is reduced when DVB is present because the required contrast for the same visual task is increased by the presence of a veiling luminance (L_v), which is the sum of the DVB contributions from all of the luminaires in the visual field of the driver (Equation 9). The coefficient 10 in this equation corresponds to an average value for 60 to 65-year-old people (9); it would be much smaller for younger people.

The effective RCS is

$$RCS^* = 1.074L/(L + L_v) \times 13.7\left\{[(L + L_v)/1.074] - 0.06\right\}^{1/2} \quad (14)$$

where $L_v = \sum_{i=1}^{n} (DVB)_i$.

The alternative possible design criterion, that of limiting the visual threshold increment (TI) (9), can be combined with the RCS standard into one diagram as shown in Figure 8. For example, the luminance values of acceptable lighting installations will be below the shaded line in Figure 8 if the requirements are RCS* ≥ 10 percent and TI ≤ 30 percent. This line is tentatively proposed as a standard for major highways and expressways that justify fixed lighting. The only additional specifications would be those of minimum point values of luminance on the traveled road surface and of illuminance on the edge of a paved shoulder.

Figure 9 presents a typical design. The performance parameters calculated for three possible lighting configurations [CL.M. = 700-W, clear mercury-vapor (type III, medium-distribution, cut off) lamps spaced 73.2 m (240 ft) apart; H.P.S. = 400-W, light-pressure sodium-vapor (type III, medium-distribution, cut off) lamps spaced 88.4 m (290 ft) apart; and L.P.S. = 180-W, low-pressure sodium-vapor (type IV, medium-distribution, cut off) lamps spaced 70.1 m (230 ft) apart] are given below (1 km = 0.6 mile, 1 lx = 0.093 ft•c, and 1 cd/m^2 = 0.29 ft•L).

	Lighting Configuration		
Performance Parameter	CL. M.	H.P.S.	L.P.S.
Avg illuminance on roadway, lx	11.6	14.3	11.85
Min illuminance on outer edge of shoulder, lx	3.7	4.3	5.6
Avg luminance on roadway, cd/m^2	0.76	0.90	0.74
Min luminance on roadway, cd/m^2	0.295	0.28	0.27
DVB (inner lane), cd/m^2	0.23	0.32	0.18
Relative energy consumption per km, W	19 140	9050	5140

The veiling luminance percentages for the three configurations are given in Figure 8. All of them are well below the shaded line. The values of the average illuminance or average luminance should not by themselves be regarded as critical.

The dimensions, such as spacing or mounting height, of the design layout should also be varied in the modeling calculations to optimize performance parameters.

COMPARISON WITH FIELD MEASUREMENTS

The Illum 1 program was used to simulate the performance of a test area of the Toronto Bypass (10). The input data used the standard reflectance surface given in Table 1, q_0 = 0.07 cd/m^2 (0.029 ft•L), which is representative of moderately old black asphalt surfaces having good skid resistance, and an estimated maintenance factor of 0.8. The values calculated were in close agreement with those measured in the field (Figure 10). This comparison is more valuable in respect to the shape of the curves than to the actual magnitudes of the luminance and illuminance because of uncertainties in the initial lamp ratings and the maintenance factor. The actual installation represents practical field conditions without very accurate alignment of luminaires.

REFERENCES

1. American National Standard Practice for Roadway Lighting. Illuminating Engineering Society, 1972.
2. A. Ketvirtis. Highway Lighting Engineering. Foundation of Canada Engineering Corp., Toronto, 1967.
3. W. Adrian. The Principles of Disability and Discomfort Glare. Visibility Symposium, Texas A&M Univ., College Station, 1968.
4. Calculation and Measurement of Luminance and Illuminance in Road Lighting. Commission Internationale de l'eclairage, Paris, Publ. 30 (TC-4.6), 1976.
5. E. Fredriksen. Lighting Quality Under Changing Weather Conditions. International Lighting Review, Vol. 23, No. 1, 1972, pp. 14-16.
6. L. E. King. Measurement of Directional Reflectance of Pavement Surfaces and Development of Computer Techniques for Calculating Luminance. Journal of the Illuminating Engineering Society, Vol. 5, No. 2, Jan. 1976, pp. 118-126.
7. A. Erbay. Atlas of the Reflection Properties of Road Surfaces. Institut fur Lichttechnik der Technischen, Univ. of Berlin, 1974.
8. A United Framework of Methods for Evaluating Visual Performance Aspects for Lighting. Commission Internationale de l'eclairage, Paris, Publ. 19 (TC 3.1), 1972.
9. W. Adrian and D. A. Schreuder. A Simple Method for the Appraisal of Glare in Street Lighting. Lighting Research and Technology, Vol. 12, No. 2, 1970, pp. 61-73.
10. C. Blamey. Photometric Measurements and Computer Simulation of the Lighting on the Toronto Bypass of Highway 401. Research and Development Division, Ontario Ministry of Transportation and Communications, Downsview, IR63, March 1977.

Publication of this paper sponsored by Committee on Visibility.

Limitation of Disability Glare in Roadway Lighting

F. W. Jung, Research and Development Division, Ontario Ministry of Transportation and Communications

Safety and comfort while driving at night depend on the visual detection of objects, which is based on contrast. The performance of this visual task is related to the relative contrast sensitivity of the lighting system provided, which is a function of roadway or background luminance and is adversely affected by disability veiling brightness or glare. The limitation of disability glare from luminaires by specifying a minimum value of effective relative contrast sensitivity for a particular road class is proposed. A simple formula has been derived for the effective relative contrast sensitivity of a lighting system by using curve-fitted, standardized data. Glare control by limiting the relative contrast sensitivity can be achieved by a permissible glare formula or a diagram. The method is demonstrated by examples. Relative-contrast-sensitivity glare control can also be combined with a method that is based on limiting the threshold increment of a critical-size object. The relative-contrast-sensitivity method and the visibility-index method both use the same concept of contrast-sensitivity change with glare.

Driving at night presents the driver with a visual task that is less comfortable and more critical for his or her safety than driving in daytime. Depending on the speed, headlights or fixed highway lighting or both are needed. If there is highway lighting, it must be of sufficient quality, since the visual task requires a certain level of luminance of the roadway surface, a certain degree of uniformity of this luminance, and a restriction of the glare from the luminaires in the visual field of the driver's eyes and from the headlights of opposing cars.

Glare has two effects on the driver: It creates feelings of discomfort and it interferes with vision. These effects are referred to as discomfort and disability glare respectively and are treated differently. In the design of roadway lighting, it is necessary to define and restrict both kinds of glare.

This report offers a concept that can be used to restrict disability glare. A driver's vision at night is concerned more with the detection than with the identification of objects. The detection of objects depends on the ability to distinguish luminance differences, and this ability is related to the reciprocal value of the contrast, which is called the contrast sensitivity. Any particular visual task requires a certain level of contrast sensitivity; the more difficult the task, the higher the level of contrast sensitivity required for it.

Methods for the evaluation of the visual-performance aspects of lighting have been recommended by the International Commission on Illumination (CIE) (2). The interference with vision by disability glare can be formulated within this framework by using well-known methods of glare calculation. This approach will develop a simple tool for evaluating the relative performance of lighting designs in terms of their relative contrast sensitivity (RCS).

RELATIVE CONTRAST SENSITIVITY

Vision during the operation of a vehicle at night is primarily a matter of the contrast between the object of vision and its background. In any visual task at low or medium luminance, if more light is available less contrast will be needed to fulfill the task successfully; i.e., higher contrast sensitivities can be permitted for the same visual task performed under better lighting.

The tabulated values of RCS as a function of background or adaptation luminance (for the vision of small objects) set the luminance of 10 000 cd/m^2, which represents daylight, at 100 percent. For lower values of background luminance, the RCS needed for the same visual task is a fraction of this maximum value.

For purposes of roadway lighting, only the lower part of this RCS function (that in the range of 0.2 to 2.5-cd/m^2 background luminance) is of interest.

As shown in Figure 1, in the luminance range between 0.15 and 2.5 cd/m^2, the standard RCS can be calculated by the following equation, which was obtained by curve fitting of the corresponding tabulated CIE values:

$$RCS = 13.7(L - 0.06)^{1/2} \qquad (1)$$

where L = luminance in cd/m^2. This standard RCS will be modified by disability glare and eye adaptation when there are glare sources present in the visual field of the driver, so that the effective RCS values (RCS_{eff}) are smaller than the standard values for the same roadway luminance.

Therefore, this concept of RCS is appropriate for combining the following requirements for the visual task of night driving: (a) a sufficient level of roadway luminance and (b) a restriction of disability glare. The combination of these two can be achieved by specifying minimum values of RCS (RCS percentages) for various classes of highway lighting installations. This approach avoids introducing specific values of contrast into standard practice at the design level.

DISABILITY VEILING BRIGHTNESS

The physiological effect of disability glare has been described in terms of a scattering of light in the eye of the driver. The amount of scattered light is larger for older people, and the effect can be calculated in terms of a disability veiling brightness (DVB) that, in a manner similar to that of a veil, reduces the contrast of night vision. Of the many formulas that have been derived to calculate this veiling luminance or DVB value, the Holladay equation (4) may be the best:

$$DVB = 10E_v/2 \qquad (2)$$

where E_v = vertical illuminance at the eye (in lux) = angle between normal line of sight (horizontal) and the glare source (in degrees). (The coefficient 10 is kept in accordance with the original reference.)

The contributions of DVB from all light sources in the driver's field of view are cumulative. Their geometric relationships are shown in Figure 4 and described by Equations 4 to 9 of the preceding paper.

Figure 2 presents a comparison of DVB equations derived by various authors (5, 6, 7, 8, 9) for a row of luminaires situated vertically above the driver's line of vision. The behavior of the function f for the very important smaller values of θ varies considerably among the different expressions, but Equation 2 appears to be as good as any other.

Because of the windshield framing, the angle (θ) is limited to the field of vision beneath an approximately 20° plane with the roadway level and therefore must be less than, or equal to, 20° (1).

Since all of the luminaires within this visual field contribute to the DVB,

$$L_v = DVB = \sum_{i=1}^{n} 10E_v/2 \qquad (3)$$

where

n = number of luminaires within the visual field of the driver, indexed with i, and
L_v = veiling luminance in cd/m².

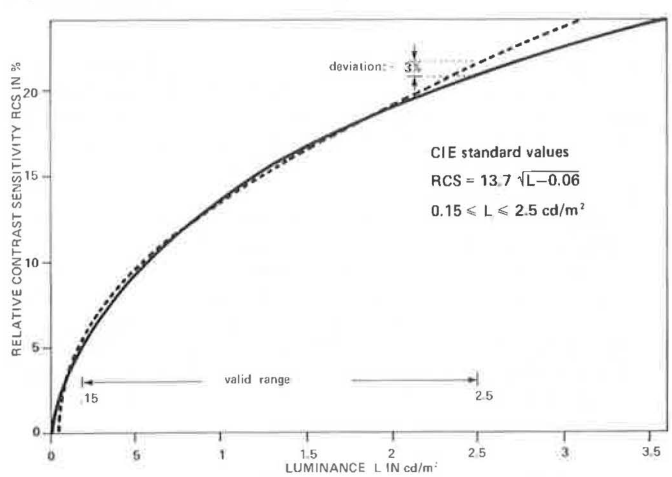


Figure 1. Relative contrast sensitivity.

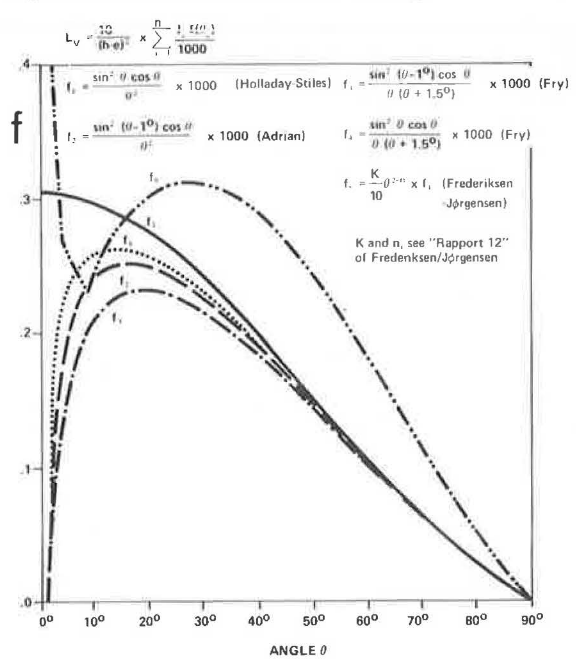


Figure 2. Functions $f(\theta)$ for C_o line or $\delta = 90°$.

PERMISSIBLE DISABILITY GLARE

The calculated luminance cannot be directly related to the night-driving task without modification. Whereas discomfort glare does not necessarily interfere with this task, disability glare does so by reducing the visibility of objects. This reduction can best be described by a reduction in contrast.

Denote

L = background luminance,
L_0 = object luminance,
L_v = veiling luminance (DVB),
C = contrast without the presence of veiling luminance, and
C_v = reduced contrast when veiling luminance is present.

Then, for small objects,

$$C = (L_0 - L)/L \qquad (4)$$

If L_v is added to both L_0 and L, this becomes

$$C_v = [(L_0 + L_v) - (L + L_v)]/(L + L_v) = (L_0 - L)/(L + L_v) \qquad (5)$$

Since the luminance difference $(L_0 - L)$ is constant, the reduction in contrast is

$$C_v/C = L/(L + L_v) \qquad (6)$$

This almost constitutes a disability-glare factor (DGF), except that, because this reduction of contrast is partially countered by human-eye adaptation (2), it must be modified. The modified factor can be written as follows:

$$DGF = [1.074L/(L + L_v)]\ RCS[(L + L_v)/1.074]/RCS(L) \qquad (7)$$

where RCS(L) = standard relative contrast sensitivity for luminance L.

When the DGF is multiplied by the standard RCS(L),

$$RCS_{eff} = DGF \times RCS(L) = [1.074L/(L + L_v)]\ RCS[(L + L_v)/1.074] \qquad (8)$$

where RCS_{eff} = modified relative contrast sensitivity that takes disability glare into account.

As illustrated in Figure 1, the standard RCS can be approximated by Equation 1, so that

$$RCS(L + L_v)/1.074 = 13.7\{[(L + L_v)/1.074] - 0.6\}^{1/2} \qquad (9)$$

Equation 9 must be substituted into Equation 8 and is valid for $0.15 \le L \le 2.5$ cd/m², which is the range of streetlighting luminance for average and minimum values. Within this range, the accuracy of the approximation is within 3 percent.

The performance of a lighting installation can be evaluated by setting permissible values for the disability glare; i.e., the RCS_{eff} provided by the lighting installation in terms of average roadway luminance under the influence of glare must not fall below a certain specified minimum value. In other words, the roadway luminance must not be so low, nor the glare so severe, that the required contrast for the visual task of night driving becomes excessively large, i.e., the contrast sensitivity must not be excessively low.

With this concept, the permissible DVB value (ΣDVB) can be established as a function of the prevailing average background luminance and an effective standard minimum value of RCS, which is specified in accordance with road classes or lighting warrants.

Denote

RCS(L) = reference relative contrast sensitivity for the average luminance of the particular lighting system (in percent),
L_{vall} = maximum permissible DVB (ΣDVB) (in cd/m²),

L = average prevailing roadway luminance (in cd/m^2),
L_v = maximum veiling luminance or DVB of the lighting system [dynamic average value for the worst 10 m (30 ft) along the lane with the most glare] (in cd/m^2), and
RCS* = minimum specified RCS for a particular road class (minimum effective value required in percent).

Then, from Equations 8 and 9

$$RCS^* = [1.074L/(L + L_{vall})]\ RCS[(L + L_{vall})/1.074] \tag{10}$$

If Equation 9 is substituted for $RCS[(L + L_{vall})/1.074]$ and Equation 10 is solved for L, a functional relation between L and L_{vall}/L is obtained:

$$L = (RCS^*/13.7)^2\ [(L + L_{vall})/1.074L] + 0.6/[(L + L_{vall})/1.074L] \tag{11}$$

or, if Equation 10 is solved for $L_{vall/L}$,

$$L_{vall}/L = 0.537KL\left\{1 + [1 - (0.24/KL^2)]^{1/2}\right\} - 1 \tag{12}$$

where

$K = (13.7/RCS^*)^2$ and
L_{vall}/L = disability glare as the allowable fraction of average roadway luminance.

Within its range of validity ($0.15 \le L \le 2.50$ cd/m^2), Equation 12 can be used to calculate a permissible maximum limit for the disability glare, $\Sigma DVB = L_v$. Figure 8 of the preceding paper shows the permissible L_v values, as a percentage of L, versus L, for various RCS* percentages. The calculated DVB (L_v) must be smaller than the permissible value that was calculated by Equation 12.

It has been suggested that the limitation of disability glare should be based on the maximum value of the calculated L_v, but this value is very sensitive to the windshield cutoff angle and the density of the selected grid points. A more appropriate base may be the average of the 10 m of driving lane that has the worst glare, which corresponds to about 11 percent of the implied stopping distance of 90 m.

COMPARISON WITH THRESHOLD-INCREMENT CRITERION

Adrian and Schreuder (9) have proposed limiting disability glare by limiting the threshold increment (TI) for the detection of a critical object within 8 min of viewing angle. For the RCS criterion, the TI due to glare is unimportant so long as the ability to detect by contrast remains at the same level for a particular visual task of night driving. For higher, effective average-luminance values of the system, the permissible TI due to disability glare may also be higher, i.e., higher average luminance may be permitted to counter the visual disability from higher glare. On the other hand, the TI criterion implies that the increase in this threshold should be about the same for installations with high, average, or low luminance.

The TI due to glare is defined by

$$TI = (\Delta L_G - \Delta L_0)/\Delta L_0 \times 100\% \tag{13}$$

where

ΔL_0 = threshold of luminance difference for detecting a critical object without glare interference and
ΔL_G = threshold of luminance difference for detecting a critical object under the influence of glare.

For given standard values of TI, L_{vall} can be calculated by using the following equation, which is valid for $0.05 \le L \le 5$ cd/m^2:

$$L_{vall} = (TI \times L^{0.8})/65 \tag{14}$$

Curves for Equation 14 are plotted in Figure 8 of the preceding paper.

For L = 0.78 cd/m^2 (which is close to a proposed standard minimum value), both criteria, TI and RCS, are identical for RCS* = 10 percent and TI = 30 percent, but this TI value is twice as high as the maximum proposed by Adrian (9).

CONCEPT OF DISABILITY-GLARE CONTROL

Gallagher (3) has carried out visibility studies that use the same concepts (2) of RCS and the visibility-glare factor presented and defined in this paper. The relation between RCS_{eff} and Gallagher's visibility index (VI) can be expressed as follows:

$$RCS_{eff} = VI/C \tag{15}$$

where C = physical contrast.

The evaluation of roadway lighting systems should be based on the RCS as defined by Equations 8 and 15, which eliminates any reference to a specified standard target for average design work. In particular, disability-glare control should be based on a diagram such as that shown in Figure 3, which shows limit lines for disability glare that, for any installation, can be plotted as percentages of average roadway luminances.

The steep limit lines at the left of the figure show that installations that have low glare can also have slightly lower values of average roadway luminance for the same level of visibility.

The horizontal limit lines are determined by approximately constant TI values. This diagram can be used directly for relative comparisons of installations. It is also potentially useful for extending this concept toward including (adding) values of headlight glare.

The RCS_{eff} or RCS* values are strictly design values to be calculated or measured by using lighting installation data only. They are not directly related to the visibility of a particular object at a particular stopping distance. This is an advantage over the visibility-index concept for average design work.

EXAMPLES

Example 1

Figure 4 (10) presents a design for a lighting installation. The average luminance and DVB values calculated for this installation with the Illum 1 program (11), which is discussed by Jung and Blamey in the preceding paper, are L = 0.686 and 0.980 cd/m^2 for black asphalt and for concrete respectively and L_v = 0.281 cd/m^2. For black asphalt at point 1A, $100L_v/L$ = 41 percent, and for concrete at point 1C, $100L_v/L$ = 29 percent.

Example 2

This example is based on an Illum 1 computer simulation of the existing lighting on Highway 401, the Toronto Bypass, and uses a maintenance factor of 0.8. The average luminance and DVB values are L = 0.68 and 0.97 cd/m^2 for black asphalt and concrete respectively and

Figure 3. Glare-control concept.

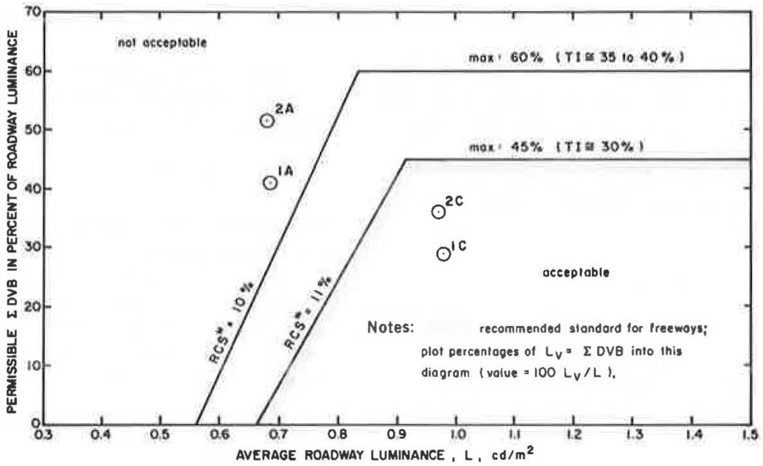
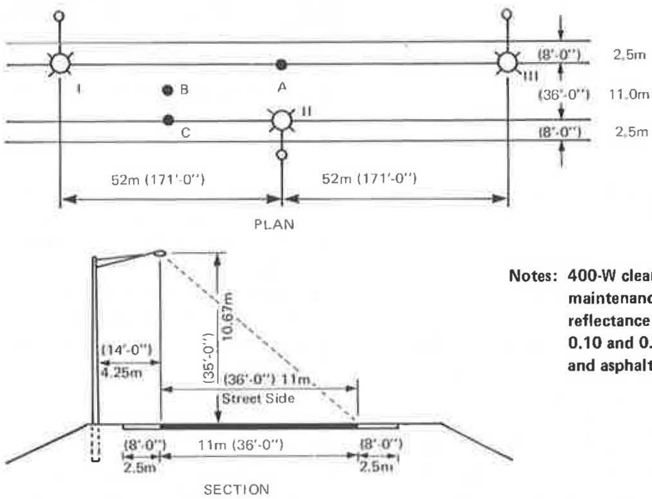


Figure 4. Design for a lighting installation.



Notes: 400-W clear mercury luminaire; maintenance factor = 0.68; reflectance surface: R3; q_o = 0.10 and 0.07 for concrete and asphalt respectively.

$L_v = 0.35$ cd/m². For black asphalt at point 2A, $100L_v/L$ = 51.5 percent, and for concrete at point 2C, $100L_v/L$ = 36 percent. The calculated disability glare is acceptable for concrete surfaces but not for black asphalt surfaces.

Neither of these examples would pass the visual TI criterion proposed by CIE, which may be too severe for North American luminaires. The maximum permissible value of TI, according to Adrian (9), is 15 percent, which is below the lower line on Figure 8 of the preceding paper.

CONCLUSIONS AND RECOMMENDATIONS

Disability glare, which is usually measured or calculated in terms of the DVB, reduces the necessary RCS_{eff} of roadway lighting for the particular visual task of night driving. The data available on the RCS (2) can be used to compare various lighting installations in terms of their visibility conditions under the influence of disability glare by establishing the percentage ratio of disability glare (ΣDVB) in relation to the average roadway luminance and then plotting these values into a glare evaluation diagram.

Any lighting design in which luminance method and disability-glare calculations have been applied can be represented by a point plotted in Figure 3, which can be used by designers to compare a variety of designs. These limit lines could also be more firmly established by research using standard targets and correlated driver-reaction time.

Recommendations about the calculation and measurement of DVB are as follows:

1. Use the Holladay formula (Equation 2) for calculations and a corresponding glare lens for measurement.
2. Calculate an average value of DVB over the worst 10 m of the worst lane, i.e., over those 10 m where the DVB is largest. This corresponds to a reasonably small fraction of the viewing or stopping distance.
3. Since the DVB depends on the visual field of the driver, and the windshield edge and the driver's eye form a limiting plane at a 20° angle with the roadway plane, evaluate this angle for passenger cars and drivers.

Since the RCS_{eff}, as defined and calculated in this paper, is identical with Gallagher's visibility index divided by the contrast (C) of his standard target, his method can be used to determine or verify standard RCS values.

REFERENCES

1. Calculation and Measurement of Luminance and Illuminance in Road Lighting. Commission Internationale de l'eclairage, Paris, Publ. 30 (TC-4.6), 1975.

2. A Unified Framework of Methods for Evaluating Visual-Performance Aspects for Lighting. Commission Internationale de l'eclairage, Publ. 19, 1972.
3. V. P. Gallagher. A Visibility Metric for Safe Lighting of City Streets. Journal of the Illuminating Engineering Society, Jan. 1976.
4. American National Standard Practice for Roadway Lighting. Illuminating Engineering Society, New York, 1972. Appendix C.
5. L. L. Holladay. The Fundamentals of Glare and Visibility. Journal of the Optical Society of America, Vol. 12, 1926.
6. G. A. Fry. A Reevaluation of the Scattering Theory of Glare. Illuminating Engineering, Vol. 49, 1954, p. 102.
7. G. A. Fry. Physiological Bases of Disability Glare. Proc., Commission Internationale de l'eclairage, Zurich, 1955, Vol. 1, 1.42.
8. W. Adrian. The Principles of Disability and Discomfort Glare. Visibility Symposium, Texas A&M Univ., College Station, 1968.
9. W. Adrian and D. A. Schreuder. A Simple Method for the Appraisal of Glare in Street Lighting. Lighting Research and Technology, Vol. 2, No. 2, 1970, pp. 61-73.
10. A. Ketvirtis. Highway Lighting Engineering. Foundation of Canada Engineering Corp., Toronto, 1967.
11. F. W. Jung and others. Computer Program for Fixed Highway Lighting. Ontario Ministry of Transportation and Communications, Interim Rept. IR 59, March 1976.

Publication of this paper sponsored by Committee on Visibility.

Effect of Improved Illumination on Traffic Operations: I-76 Underpass in Philadelphia

Michael S. Janoff, Franklin Institute Research Laboratories, Philadelphia

An experimental lighting system in an underpass on I-76 in Philadelphia was evaluated. The lighting system was designed to provide five levels of illumination ranging from 5382 lx (500 ft·c) horizontal to 22 lx (2 ft·c) horizontal. Low-pressure sodium-vapor lamps were used. The internal level was automatically set by a series of photocells external to the underpass and provided a ratio of internal to external illuminance of approximately 10 percent. Four measures were used to determine the effect of the improved illumination on traffic operations. These were (a) the effect on the number of traffic accidents, (b) the effect on vehicle-velocity maintenance, (c) the effect on deceleration (braking) characteristics, and (d) the effect on subjective responses of drivers to the new lighting. The photometric characteristics of the new lighting were evaluated and the Illuminating Engineering Society and the American Association of State Highway and Transportation Officials tunnel-lighting recommendations were compared. The results indicated that (a) the new lighting caused decreases in the velocity variability and in brake applications at the portal, (b) in general, as the internal lighting level increased, both the velocity variability and the number of brake applications decreased, indicating safer and smoother traffic operations, (c) drivers responded positively when the internal lighting levels were increased and there were no noticeable adverse effects caused by the low-pressure sodium-vapor lamps, (d) the Illuminating Engineering Society recommendations for tunnel lighting appear to be preferable to those of the American Association of State Highway and Transportation Officials, and (e) there was a reduction in the number of accidents inside the underpass and at the portal in the 6 months after installation of the new lighting.

The first object of this program was to evaluate the effects on traffic operations of lighting improvements in the eastbound section of the Thirtieth Street underpass on I-76 in Philadelphia. The improvements included a variable-level lighting system, the resurfacing of the roadway, and new reflective walls (both the side wall and a temporary center wall).

The second object was to determine whether the selected luminaires are adequate for their purpose and to compare whether Illuminating Engineering Society (IES) (1) or American Association of State Highway and Transportation Officials (AASHTO) (2) recommendations are better design guidelines for tunnel lighting.

SYSTEM DESCRIPTION

Original Lighting System

The original daytime lighting system (the before condition) consisted of two rows of 1500-mA fluorescent lamps supplemented by thirteen 400-W mercury-vapor lamps in the first 73 m (240 ft) of the underpass. The illumination provided by this system during the daytime was approximately 355 lx (33 ft·c) at an average position and about 538 lx (50 ft·c) at the portal entrance.

Present Lighting System

The present lighting system (the after condition) consists of five continuous rows of overhead fixtures in the first 49 m (160 ft) of the underpass, one row of fixtures in the next 30 m (100 ft), and the original fluorescent lamps for the remainder of the tunnel. Each fixture in a row houses two 180-W low-pressure sodium-vapor lamps, except that, in the center row, a 90-W lamp is substituted for one of the 180-W lamps in every eighth fixture.

The electrical circuitry is designed so that five different lighting configurations are possible. The control is monitored by a series of four photoelectric cells mounted outside the underpass. The inside design levels, the outside illuminations at which the circuits are energized, and the configurations are summarized below (1 lx = 0.093 ft·c).

Circuit	Illumination (lx) Inside	Outside	Configuration
N1	27	Night	All 90-W lamps, one in every eighth fixture in center
D1B	538	54	Center row, one lamp in every fixture
D1A	1076	5 382	Center row, two lamps in each fixture
D11	3220	21 529	D1A plus one lamp in each fixture of remaining four rows
D21	5382	43 057	All lamps

PROCEDURES AND RESULTS

Experimental Method

The following interrelated experiments were designed to evaluate the effectiveness of the new lighting and to compare the IES and AASHTO recommendations:

1. A measure of driver performance in terms of individual speed profiles near the tunnel portal,
2. A measure of driver performance in terms of the number of brake applications near the tunnel portal,
3. A survey of accident histories near the tunnel portal,
4. A photometric measure of the illumination (and luminance) of the new luminaires, and
5. A survey of subjective driver responses to the new lighting system.

Driver Performance in Terms of Individual Speed Profiles

All of the velocity-profile data were collected by using the tape-switch system designed by the Franklin Institute Research Laboratories (3). The records of velocity variability (with the temporary center wall in place) as indicated by the individual standard deviation for each measured vehicle were grouped into six groups stepped in one-half sigma units, and χ^2 tests of significance, based on the unique independent variable (clear, cloudy, D1A, D11, D21, and nighttime), were made on each of these matrices.

A review of the raw variability data indicated that the velocity maintenance was least variable at night. This is presumably because of the relatively low visual difficulty or of a more stable visual-difficulty level in the transition from the exterior to the interior of the underpass and suggests that the nighttime driver behavior represents the optimum case of velocity maintenance. The statistical comparison of the velocity variability for the before and after conditions versus nighttime driver performance (shown below) demonstrates the relative success of each of the lighting alternatives at achieving this minimum variability level.

Condition	Probability Versus Nighttime	Condition	Probability Versus Nighttime
Before		After (clear day)	
Clear	0.005	D1A circuit	0.005
Cloudy	0.25	D11 circuit	0.10
		D21 circuit	0.25

The nighttime case, when compared to the clear case (unmodified lighting on a bright day) is statistically significant to a convincing degree. The D1A case clearly indicates the inadequacy of this system, and the progressively lower significance levels of the other alternatives indicate the increasing visual quality that the higher lighting level represents.

Figures 1 through 5 illustrate the speed profiles in the various before and after lighting conditions. Figure 1 illustrates the mean and 85th percentile speed profiles for clear, cloudy, and nighttime conditions for the before lighting condition. Figure 2 compares the nighttime (optimal) condition with each of the three after lighting conditions during clear weather. (The higher velocity for the nighttime condition is attributed to lower traffic volumes; only the change or variability in velocity is of significance here.) Figures 3, 4, and 5 show that, for the D1A case, there is a significant decrease in velocity as vehicles approach the entrance to the underpass and that this decrease is substantially reduced in the D11 and D21 conditions. The after cloudy condition shows no significant differences between the nighttime (optimal) condition and the D11, D1A, and D1B conditions.

The records of velocity variability after removal of the temporary center wall showed no significant differences among any of the after lighting conditions (D1A, D11, or D21). In comparison with the before cases in which nighttime is the optimum and a clear day is the worst, the D11 after condition has the least velocity variability and is closest to the optimum nighttime-before case. The D1A case shows a decrease in velocity at about 30.5 m (100 ft) inside the portal. The D21 case shows a decrease in velocity before the tunnel portal that is maintained for at least 61 m (200 ft) (the limit of the recording equipment). None of the differences was significant. The results are summarized in Figure 6.

Driver Performance in Terms of Brake Applications

With the hypothesis that, as the internal lighting was increased, the frequency of braking, which indicates driver uncertainty, would decrease, an observer was stationed downstream about 183 m (600 ft), with clear visibility into the southbound entrance. The brake lights that appeared at or directly inside the tunnel entrance were counted (unless the light was due to the braking of a lead automobile), and the horizontal illumination and the traffic volume in the southbound lanes were measured continuously (every half hour). Figure 7 summarizes the results of this experiment (with the center wall in place). As the lighting in the underpass increased under a relatively constant outside illumination, the number of drivers activating their brakes at or near the entrance to the underpass steadily decreased, which indicates that, as the interior lighting levels decrease, the decreasing visibility causes driver uncertainty and a greater tendency to brake near the tunnel entrance. The same effect was observed after removal of the center wall and is shown in Figure 8.

Accident Analysis

The before-condition accident data were obtained from the Pennsylvania Department of Transportation (PennDOT) for the period of 1969 to 1972 and from the city of Philadelphia for the period of 1968 to mid-1973. The after-condition accident data were available only from the city of Philadelphia for the period of June to November 1974. Any conclusion drawn from these results will be considered as tentative. These data are summarized below.

Location	Accidents: Before Condition Total	No./Year	After Condition Total	No./Year
Entrance ramp, tangent section, and underpass	194	35	16	32
Tangent section and underpass	156	28	10	20
Underpass only	108	20	4	8

There are not sufficient data for further meaningful stratification, but these comparisons seem to indicate a reduction in the number of accidents.

Photometric Measure of New Luminaires

The illumination provided by the before-condition lighting system was measured by using automatic recording equipment (4); the results are shown below.

Condition		Illumination (lx)			
		Horizontal		Vertical	
External	Internal	Avg	Avg/Min	Avg	Avg/Min
Day	Day	355	3.0	–	–
Night	Day	237	2.0	92	1.7
Night	Night	58	11.0	–	–

(During the daytime there is a significant amount of illumination provided by the sunlight entering from the sides, so that the average during the day is higher than that during the night for the same lighting configuration.)

The same procedures and the same equipment were used to measure the illumination provided by the new lighting system; these results are shown below.

Internal Condition	Design Level	Illumination (lx)			
		Horizontal		Vertical	
		Avg	Avg/Min	Avg	Avg/Min
N1 (night)		26	4.0	22	5.1
D1B	538	1195	–	377	–
D1A	1076	2164	–	721	–
D11	3229	5167	–	1776	–
D21	5382	7427	–	2992	–

Figure 1. Velocity versus distance from portal for three before lighting conditions.

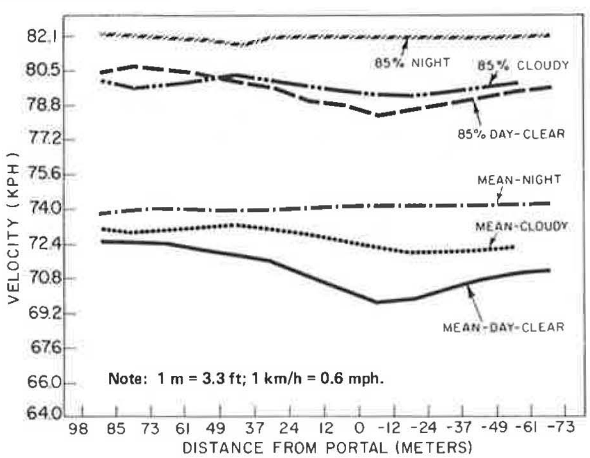


Figure 2. Velocity versus distance from portal for three after lighting conditions.

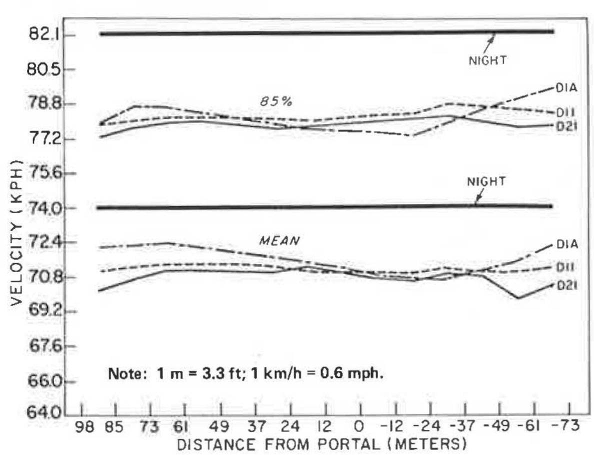


Figure 3. Velocity versus distance from portal for one before and one after lighting condition (D1A case).

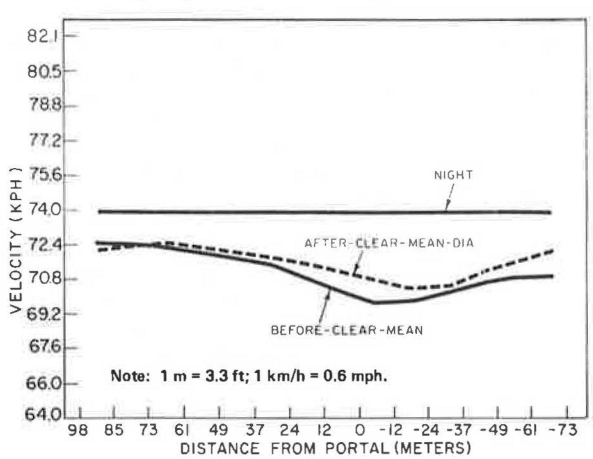


Figure 4. Velocity versus distance from portal for one before and one after lighting condition (D11 case).

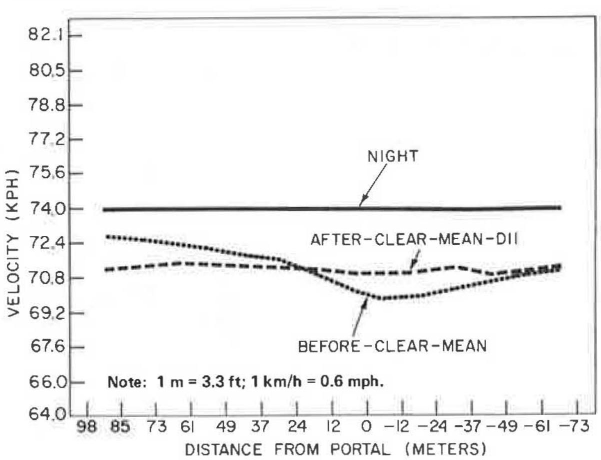


Figure 5. Velocity versus distance from portal for one before and one after lighting condition (D21 case).

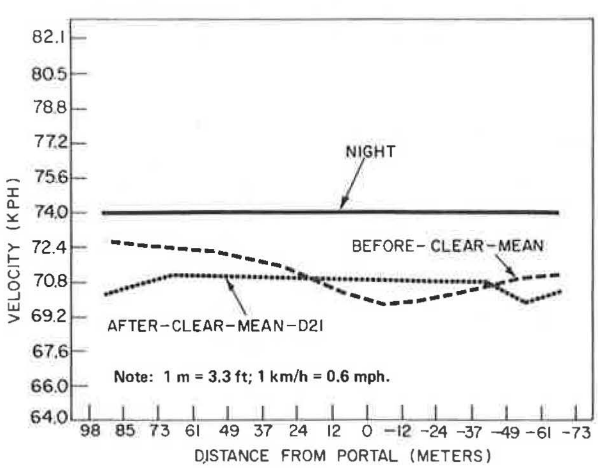


Figure 6. Velocity versus distance from portal for three after lighting conditions.

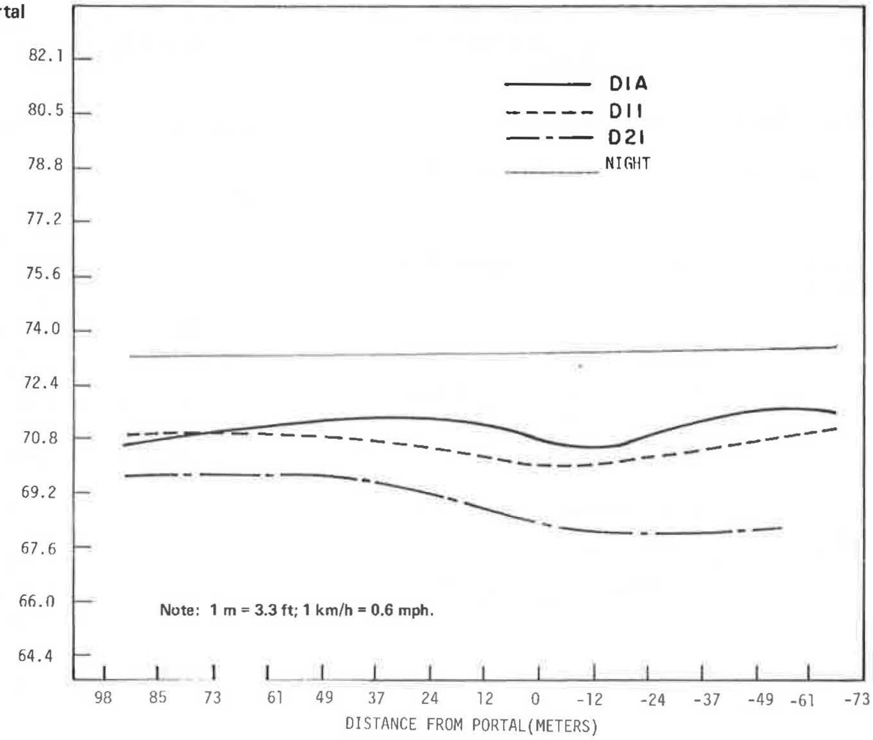
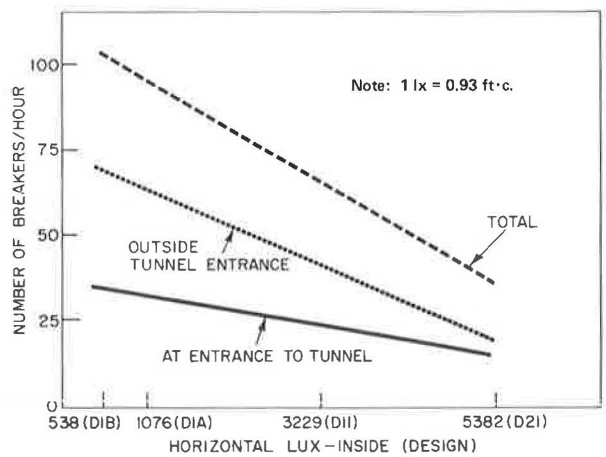


Figure 7. Number of brakers versus inside tunnel lighting (with center wall).



A complete goniometric analysis of the low-pressure sodium-vapor luminaire was also performed (5).

Driver Survey

The purpose of the survey was to determine the subjective response of the driving public to the new lighting system. Since the survey was short, required no postage by the driver, and was given to him or her immediately after the trip through the underpass, a high return rate and consistent answers were expected, but this did not occur. The response rate was low (10 percent), and discrimination between the different lighting levels was not possible. The following questions were asked.

1. How would you rate the lighting in this tunnel?
2. How was your trip through the tunnel?
3. How do you feel about this tunnel?
4. How do you feel about tunnel driving in general?
5. How can this tunnel be improved? (List in order of preference: more light, less noise, wider lanes, fewer cars, fewer trucks, cleaner, higher ceiling, or higher speeds.)
6. Do you object to using tunnels?
7. Do you have any general comments about this tunnel or tunnel driving?

Approximately 1000 mail-back surveys were distributed to motorists who had driven south through the I-76 tunnel with the center wall in place during a clear day [approximately 80 732 lx (7500 ft·c) horizontal illumination] by handing the forms to them as they exited at the South Street off-ramp. The return rate was almost equally distributed among the four daytime internal-lighting levels.

There were few data from the responses to questions 1 to 4 that could be used to discriminate statistically among the four levels of illumination. Some of this may be attributed to the fact that, in most cases, there was a lag between the time the survey was distributed and the time it was returned. Since many drivers use the underpass frequently, repeated passages under different lighting levels may have confused the results.

The responses to questions 5 and 6 were more significant. Drivers consistently chose more light and wider lanes as important variables and considered higher ceilings and higher speeds less important. Drivers who passed through the tunnel under the two lowest levels of illumination objected to using tunnels twice as often as did drivers who passed through the tunnel under the two brightest illumination levels (25 versus 12.5 percent respectively) possibly indicating a more negative attitude under lower lighting levels.

A survey of motorists who had driven through the tunnel after the removal of the center wall produced responses that were similar to those to the first survey

Figure 8. Number of brakers versus inside tunnel lighting (without center wall).

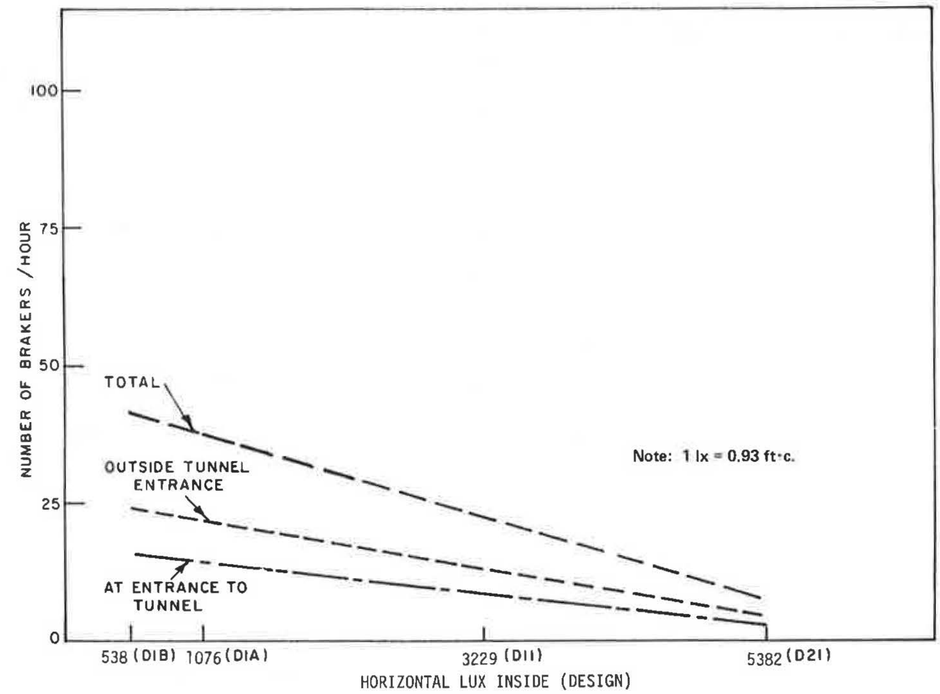
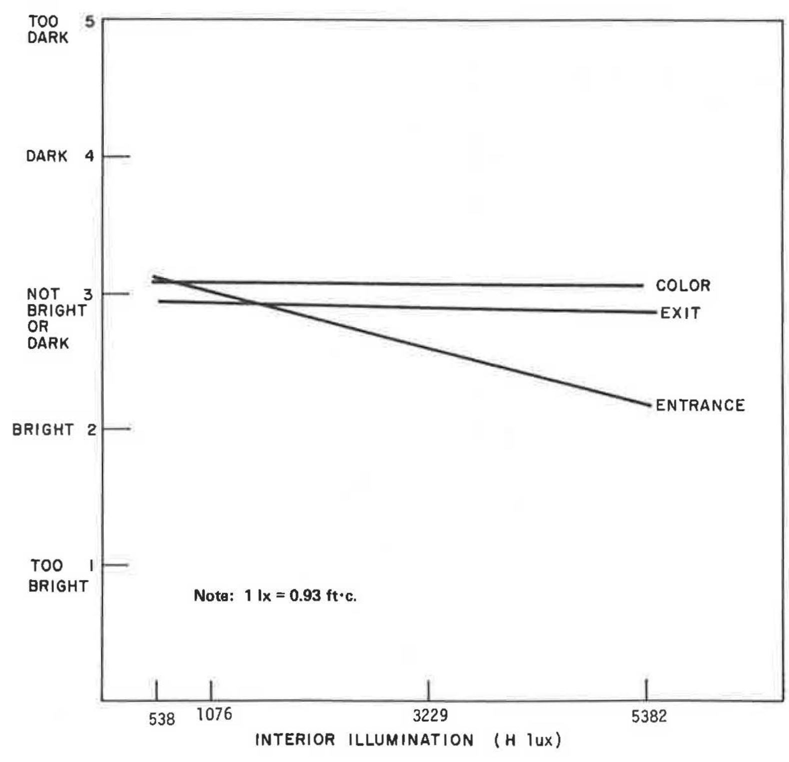


Figure 9. Subjective driver response versus interior lighting condition.



and did not differentiate among the various lighting conditions.

To supplement this survey, 84 drivers were stopped and asked question 1 after passing through the tunnel under the four different lighting conditions (D21, D11, D1A, D1B). These personal responses indicated that the effect of increasing the entrance lighting has been positive; i.e., as the level of lighting increased from D1B to D21, drivers subjectively responded that the tunnel entrance lighting appeared brighter. There was no difference in response for exit lighting, which would be expected since the exit lighting was the same for all four entrance lighting conditions. The results are illustrated in Figure 9.

COMPARATIVE EVALUATIONS

The objects were to compare the IES and AASHTO tunnel-lighting recommendations and the two after conditions (i.e., with and without the center wall in place).

Figure 10. Velocity versus distance from portal for four luminance ratios.

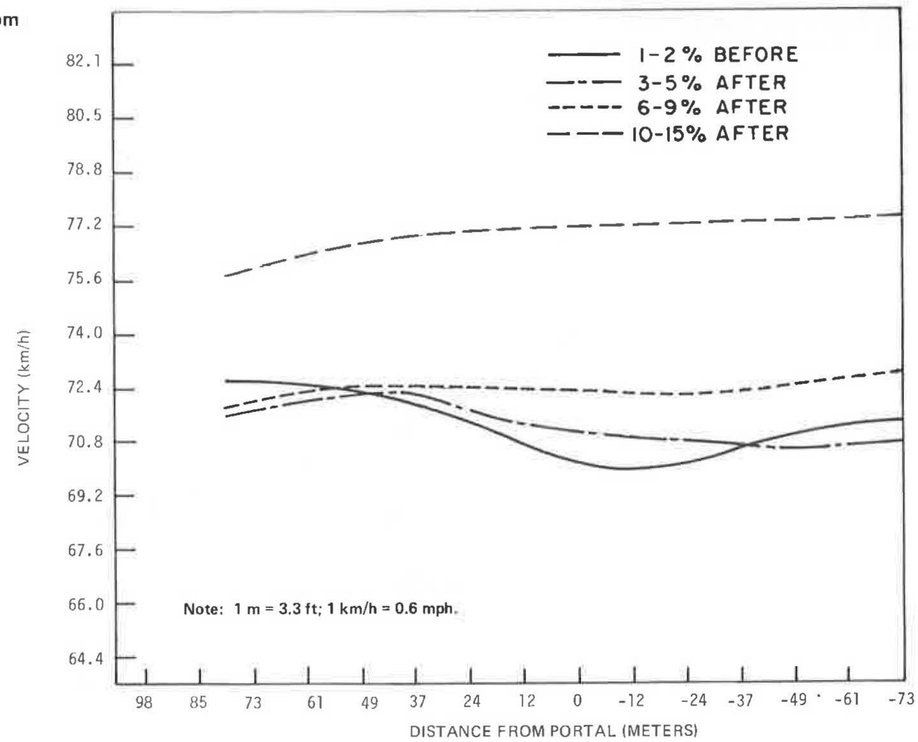
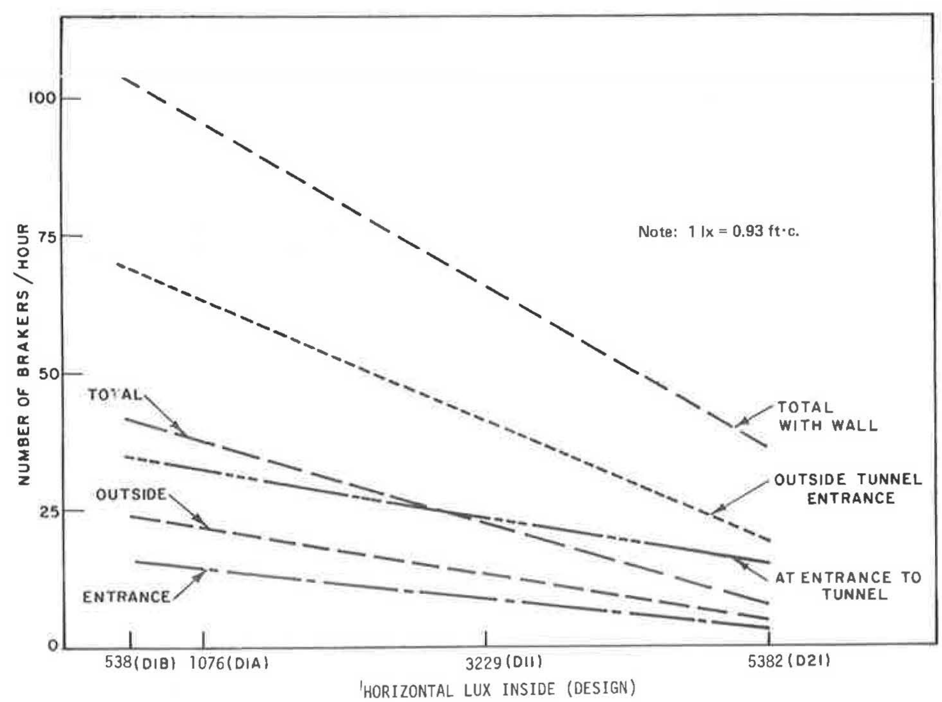


Figure 11. Summary of number of brakers versus inside tunnel lighting.



IES Versus AASHTO Recommendations

An internal horizontal illumination of about 5382 lx (500 ft•c), as provided by the D11 condition, appeared to optimize driver velocity-maintenance performance. The D21 condition provided a slightly reduced variability in the after-condition experiment with the wall in place, but the results were not significantly different from those of the D11 condition. The IES Standard recommends 5382 lx (500 ft•c) for the threshold illumination, which lasts 2 s [about 46 m (150 ft) at highway speeds] (1), but the AASHTO Standard recommends only 323 to 646 vertical lx (30 to 60 ft•c) on the tunnel wall (somewhat less than that measured for the D1B lighting condition), which was inadequate.

These standards were derived for illuminance measurements only. However, the luminances of sky, pavement, portal, and interior walls were also measured during the data collection. A preliminary evaluation of the effect of the ratio of external (sky) to internal (wall) luminance on driver performance is illustrated in Figure 10.

The before case is the worst case and has a ratio of internal to external luminance of between 1 and 2 percent. The three after cases had ratios of 3 to 5, 6 to 9, and 10 to 15 percent respectively. For the 3 to 5 percent case there is still a decrease in velocity at the portal, indicating insufficient luminance inside the tunnel, but the velocity remains stable for both the 6 to 9 and 10 to 15 percent cases. These results indicate that the ratio of internal to external luminance should be greater than 5, and probably between 6 and 9 percent.

If 7.5 percent is used as a design figure, then, for outside luminances of 17 130 to 34 260 cd/m^2 (5000 to 10 000 ft•L) (bright day conditions), the internal luminance should be 1285 to 2570 cd/m^2 (375 to 750 ft•L). These values are in the D11 to D21 lighting-system range, again indicating that the IES recommendations are better.

The results are similar for illuminance ratios. The velocity profiles at 3 to 5 percent showed a significant decrease in speed preceding the portal, those at 6 to 10 percent showed a slight reduction in speed, and those at 11 to 15 percent showed no decrease.

Effect of Center Wall

Three measures were used to evaluate the effect on driver performance of the temporary reflective center wall. These were velocity maintenance, braker data, and survey responses. The most meaningful comparisons were those from the braker data. Figure 11, which combines Figures 7 and 8, shows that the absolute number of brakers decreased after the wall was removed, although the effect of increasing the internal illumination was the same in both after cases.

The velocity measurements with the wall in place clearly showed that the D1A condition was inadequate, but with the wall removed the significant differences disappeared. This indicates a better visual field and better driver performance without the wall.

CONCLUSIONS

1. The new tunnel lighting system in the Thirtieth Street underpass on I-76 has provided a substantial improvement in visibility, velocity maintenance, and driver performance. The measured luminances (1 cd/m^2 = 0.292 ft•L) for the four lighting conditions are summarized below.

Internal Condition	Luminance (cd/m^2) Wall	Luminance (cd/m^2) Pavement (avg)
D1B	394	343
D1A	480	411
D11	1182	771
D21	1199	891

2. The optimum lighting levels for bright days are provided by the D11 system.
3. The IES tunnel-lighting recommendations appear to be more accurate than the AASHTO recommendations.
4. The effect of the center wall was not positive; driver performance was slightly better after the wall was removed.
5. The system meets or exceeds the IES lighting recommendations in terms of horizontal illumination on the pavement surface.
6. The responses of drivers to the increased light indicated an apparent awareness of it and no apparent dislike of the monochromatic lamps.
7. Six months of after-condition accident data indicate a reduction in the number of accidents inside and preceding the tunnel.

ACKNOWLEDGMENTS

This research was conducted by the Transportation Sciences Laboratory of the Franklin Institute Research Laboratories for the Bureau of Materials, Testing, and Research of the Pennsylvania Department of Transportation. The assistance of William G. Weber, Jr., of PennDOT is acknowledged.

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The United States government assumes no liability for its contents or the use thereof.

The contents of this paper reflect the views of the Franklin Institute Research Laboratories, which is responsible for the facts and the accuracy of the data presented. The contents do not necessarily reflect the official views or policies of the Federal Highway Administration, U.S. Department of Transportation. This paper does not constitute a standard, specification, or regulation.

REFERENCES

1. Lighting Handbook. Illuminating Engineering Society, Fifth Ed., 1972, pp. 20-10 to 20-14.
2. An Informational Guide for Roadway Lighting. AASHTO, March 1969, pp. 14-17.
3. E. Farber and others. Determining Pavement Skid-Resistance Requirements at Intersections and Braking Sites. NCHRP, Rept. 154, 1974.
4. V. P. Gallagher and M. S. Janoff. Interaction Between Fixed and Vehicular Illumination Systems. Franklin Institute Research Laboratories, Philadelphia; Federal Highway Administration, Nov. 1972, pp. 2-5 to 2-7.
5. M. S. Janoff. Evaluation of Experimental Tunnel Lighting in Thirtieth Street Underpass on I-76 in Philadelphia. Franklin Institute Research Laboratories, Philadelphia; Pennsylvania Department of Transportation, March 1975, Appendix A.

Publication of this paper sponsored by Committee on Visibility.

Evaluating Nighttime Sign Surrounds

H. L. Woltman and W. P. Youngblood, Minnesota Mining and Manufacturing Company, Saint Paul

The accuracy of a variety of instruments that might be suitable for field measurement of nighttime sign surrounds was evaluated by comparing measurements made with them with measurements made with a laboratory-quality telephotometer. A technique for the evaluation of surrounds that identifies them by luminance measurements was developed. The measurement of numerous surrounds leads to the conclusions that conventional descriptions are often inappropriate, that opposite sides of the same roadway may vary in luminaire intensity, and that roadway geometrics may cause variations in surrounds. Photographs and luminance values that represent four generalized luminance levels and a description of each are given.

Research on the nighttime performance of signs has shown a close relation between the luminance of a sign versus its nighttime background or surround and its visibility and legibility. However, beginning with the work of Smyth (1) and continuing to the present time, this research has used either blank laboratory surrounds that were varied in luminance only for purposes of the research or natural night surrounds that were identified only by verbal descriptions or pictures. There has been little or no systematic work that measured the highly variegated night surrounds occurring on highways (particularly those in the vicinity of official traffic signs).

Lythgoe (2), Smyth (1), Allen and Straub (3), Allen and others (4), and Forbes (5) have shown that increased sign luminance is required where sign surrounds possess increasing luminance. This is in agreement with the requirements of some standards (6, 7). Some of the luminance values for sign legends or backgrounds as a function of surround luminance are summarized below (1 cd/m^2 = 0.292 ft·L).

Investigator	Surround Luminance (cd/m^2) Dark	Medium	Bright
Smyth	15 to 25	25 to 65	65 to 170
Illuminating Engineering Society	25	50	100
Allen and others	35	70	350
Forbes	2.9 to 26	26 to 87	87 to 274

Surround luminance is high at night in urban locations where street lighting, advertising signs, and commercial lighting displays form the background for essential traffic signs. It is low on dark, rural, two-lane roads that have low traffic volumes and few intersections. Greater understanding of the spectrum represented by these extremes is desirable for

1. Accurate identification of the nighttime surround,
2. Selection of appropriate materials to achieve the necessary luminance levels and ratios of contrast, and
3. Achievement of the maximum economic benefit by the selection of materials that are appropriate to the environment of the sign.

Various federal specifications (8, 9) describe numerous performance levels for reflective materials, and there are a wide variety of lighting designs and luminaire fixtures available. However, to select the appropriate sign luminances, the nighttime surrounds should be measured and identified first. Thus, this paper evaluates practical methods of measuring nighttime sign surrounds and six available instruments, and presents a survey of measurements made with some of them.

The selection of suitable instruments must recognize the extremely varied nature of the roadside surround. Woltman (10) has reported an inventory of sign surrounds for daylight, and luminance observations of dark, rural, nighttime sign surrounds have been made (11), but the best photographs cannot convey the variety of luminance levels that occur, and the colors and extremes of contrast, both dynamic and static, to which the driver is subjected. As Luckiesh (12) points out,

> A thorough diagnosis of visibility and seeing conditions involves
>
> 1. Brightness levels of the task and the immediate and entire surroundings,
> 2. Brightness contrast between critical details and their background,
> 3. Brightness ratio of the surroundings and the task, and
> 4. Brightnesses and brightness ratios in the entire visual field.

VISUAL FIELD

The surroundings of the visual task can include the entire visual field, but there are practical reasons for limiting it. Luckiesh has noted that

> At 30° from the optical axis, visual acuity is only 1 percent of its value in the central 1° field. The effect of a glare source, and also the effect of brightness of the surroundings, decrease as the angular distance from the line of vision increases.

Matson (13) and Greenshields (14) consider the visual field of a driver within the confines of an automobile and busy with the driving task to be 6 and 10°. The majority of roadside shoulder and overhead signs are within this field.

The act of seeing fine detail is accomplished in a small field (about 1° in extent) on the optical axis of the eye. Glare sources close to this field are the most troublesome, particularly at night, when the critical task of sign reading may involve relatively low luminances and short time intervals.

The central field contains the visual task (the sign and the most important elements of the surround). According to Finch (personal communication), interfering luminances are those in a 3 to 5° field, and an average expressed as an integrated value of such sources is necessary. Olson (personal communication) agrees that an average luminance measure that surrounds the sign to the extent of one or two sign diameters is probably satisfactory.

Although other methods of evaluating the sign surround have been considered, the most immediately practical are the Pritchard type telephotometers that can selectively evaluate discreet areas of interest. The use of such instruments at sign-reading distances permits the measurement of the luminance level of the task and the surround and the determination of luminance ratios and luminances of any objects or surfaces in the visual field. A selection of probe sizes is available, and integration over an area of 1° is possible by the use of the 1° probe. This is of particular importance for the measurement of surround luminance: One degree is equivalent to a diameter of 0.53 m (21 in) at a distance of 30 m (100 ft), and at a distance of approximately 90 m (300 ft) the 1° probe gives an integrated reading of a 1.5-m (5-ft) diameter area, where the sign itself displaces approximately 0.5°. Thus, a 2.5° field, which corresponds closely to the recommended 1 to 3° field, can be examined by measuring tangentially at the edges of the sign.

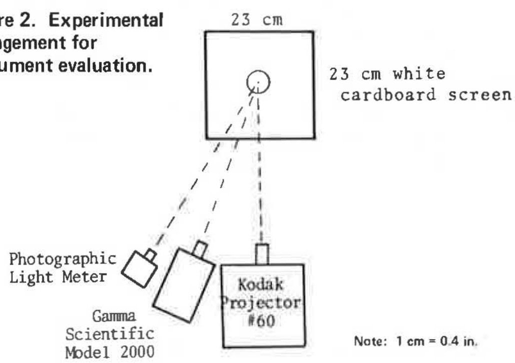
INSTRUMENT EVALUATION

The following small, portable instruments (Figure 1) that might be suitable for field measurements of surround luminances were evaluated.

Figure 1. Instruments evaluated for the measurement of nighttime surrounds.

Note: Left to right at front: instruments 2, 3, 6, 5, 4, and 3. At rear: instrument 1.

Figure 2. Experimental arrangement for instrument evaluation.



Note: 1 cm = 0.4 in.

1. Gamma Scientific, Inc., model 2000 telephotometer: This is a scientific, Pritchard type instrument and has a transistorized photomultiplier and electrometer amplifier, a portable power supply, a 1° sensing probe (an acceptable angle), photopic color correction, a measurement range of 0.003 to 120 000 cd/m² (0.001 to 35 000 ft•L), and internal standardization and calibration. It was calibrated with a National Bureau of Standards source over a number of tests and averaged ±2.5 percent.

2. Minolta TV Auto-spot: This instrument is essentially a studio, spot-reading, photographic light meter with a cadmium sulfide cell. The 1° measured area is enclosed by an illuminated etched circle. The output is given in footlamberts, and the range is 1.1 to 17 140 cd/m² (0.32 to 5000 ft•L).

3. Honeywell Pentax 1°/21° meter: This instrument is essentially a studio, spot-reading, photographic light meter with a cadmium sulfide cell. The 1° measured area is enclosed by an etched circle, which may be illuminated. The output is given as a light level with a range of 3 to 18, which corresponds to 1.0 to 34 280 cd/m² (0.3 to 10 000 ft•L).

4. Soligor Spot Sensor: This instrument is essentially a studio, spot-reading, photographic light meter with a cadmium sulfide cell. The 1° measured area is enclosed by an etched circle in the viewing field. The reticule is not illuminated. The output is given as an exposure value range of 3 to 18, which corresponds to 1.0 to 34 280 cd/m² (0.3 to 10 000 ft•L).

5. Gossen Luna-Pro: This instrument is a 30° reflected light or incident light-measuring, studio, photographic light meter. It has an incident light range of 0.17 to 344 320 lx (0.016 to 32 000 ft•c).

6. Sekonic Auto-Lumi model 86: This instrument is a 30° reflected light meter. It has an exposure value range of 6 to 18, which corresponds to approximately 6.5 to 27 425 cd/m² (1.9 to 8000 ft•L).

The experimental arrangement for the evaluation of instruments 2 through 6 relative to instrument 1 is shown in Figure 2. The photographic light meters, the scientific telephotometer, and the projector were all in the same plane. A 1° probe was used. The screen was moved from 1 to approximately 12 m (3.3 to 39 ft) to decrease the luminance.

The nighttime evaluation of these meters showed reasonably close correlations between meters 1 and 2

Table 1. Laboratory data for correlation of scientific telephotometer and photographic light meters.

	Measured Target Luminance (cd/m²)											
Instrument	White Target								Red Target		Blue Target	
No. 1												
Gamma Scientific	2.2	5.6	11.0	18.8	26.7	28.4	28.7	61.7	5.9	17.5	1.8	3.9
No. 2												
Minolta	1.4	3.1	6.9	13.7	13.7	25.7	37.7	54.8	4.3	17.5	1.5	3.8
No. 3												
Honeywell Pentax	1.7	4.8	8.6	19.9	24.0	26.7	37.7	51.4	5.5	22.3	2.1	4.5
Honeywell Pentax 2	2.1	5.8	10.3	19.2	25.7	28.4	38.7	51.4	6.9	17.5	2.7	6.9
No. 4												
Soligor Spot Sensor	2.1	3.4	8.9	16.4	20.6	22.3	32.6	46.3	5.5	17.8	1.4	2.4

Note: 1 cd/m² = 0.292 ft•L.

Table 2. Correlation of scientific telephotometer and photographic light meters.

	Error Relative to Measurement by Meter 1 (%)											
Instrument	White Target								Red Target		Blue Target	
Minolta	-37	-44	-38	-27	-49	-10	-3	-11	-27	-2	-19	-4
Honeywell Pentax	-21	-14	-22	5	-10	-6	-3	-17	-6	27	13	14
Honeywell Pentax 2	-5	5	-6	2	-4	0	0	-17	17	0	51	75
Soligor Spot Meter	-5	-38	-19	-13	-23	-22	-16	-25	-6	2	-25	-39

Table 3. Correlation of scientific telephotometer and photographic light meters.

Instrument	Avg Error Relative to Measurements by Meter 1 (%) All Targets	White Target	Red Target	Blue Target
Minolta	-22.6	-27.3	-14.5	-11.5
Honeywell Pentax	-3.3	-11.0	10.5	13.5
Honeywell Pentax 2	9.8	-3.1	8.5	63.0
Soligor	-19.1	-20.1	-2.0	-32.0

Figure 3. Correlation of scientific telephotometer and photographic light meters.

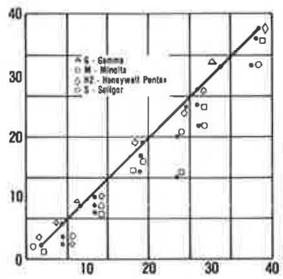
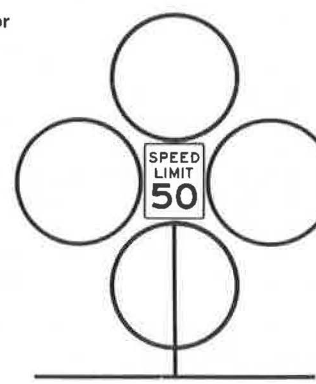


Figure 4. Probe locations for measuring sign-surround luminance.



through 4. Under high luminance conditions, meters 5 and 6 indicated higher light levels than the actual sign surrounds, and under many moderate nighttime conditions, readings lower than those measured by meters 1 through 4 were common. The wide acceptance angle of these instruments, 30°, includes too much background—either luminaires of the bright surround or black sky of the darker surround. Thus, the surround immediately adjacent to the sign was not measured as accurately as the 1° acceptance angles of instruments 1 through 4.

The laboratory data for the correlation of meters 2 through 4 relative to meter 1 are given in Table 1, and the correlation is given in Tables 2 and 3 and illustrated for the luminance range of 3.5 to 34.5 cd/m² (1.0 to 10.0 ft•L) in Figure 3. The correlation is generally linear, but the values obtained with the less expensive instruments are somewhat lower than those obtained with the laboratory instrument.

FIELD EVALUATION OF NIGHTTIME SURROUNDS

The nighttime luminance of the dark sky above, to the immediate right, below, and to the immediate left of 90 signs was measured with the laboratory instrument, as illustrated in Figure 4. The measurements were made on dark, rural sections of interstate routes in winter against earth, sky, and snow-covered backgrounds. The presence of snow or moonlight appeared to be of little significance.

The use of instruments 2, 3, and 4 involves some compromises and requires some improvements in data gathering since these instruments do not read below 0.86 cd/m² (0.25 ft•L). (Values up to approximately 2 cd/m² (0.6 ft•L) represent dark surrounds.) However, by the use of these instruments, more data can be gathered with greater convenience and less training of the operators than with the larger laboratory instrument. In practice, the driver operates the smaller instrument, and another person records the data. The output of two of the instruments is read in exposure values (EV), a

Table 4. Distribution of nighttime sign surrounds for various study sites.

Location	Dark (less than 2 cd/m²) Measured Value (cd/m²)	% of Total	Slight (2 to 6 cd/m²) Measured Value (cd/m²)	% of Total	Moderate (6 to 17 cd/m²) Measured Value (cd/m²)	% of Total	Bright (17 cd/m²) Measured Value (cd/m²)	% of Total	Total
Detroit									
Woodward Ave.	4	17	4	17	11	45	5	21	24
Grand River Ave.	0	–	2	18	3	27	6	55	11
Telegraph Rd.	0	–	5	19	16	62	5	19	26
Dearborn, Mich.									
Michigan Ave.	0	–	1	10	4	40	5	50	10
Kalamazoo, Mich.									
Michigan Ave.	0	–	5	25	11	55	4	20	20
Kalamazoo Ave.	3	18	6	35	7	41	1	6	17
Lansing, Mich.									
Saginaw St.	1	11	2	22	2	22	4	45	9
I-496	5	56	0	–	3	33	1	11	9
MI-143	0	–	2	28	0	–	5	72	7
US-127	5	62	2	25	1	13	–	–	8
Minneapolis									
Lake St.	0	–	1	6	7	47	7	47	15
I-35W	0	–	1	50	1	50	0	–	2
St. Paul									
White Bear Ave.	4	50	2	25	2	25	0	–	8
Unlighted, rural Interstate highways in Calif., Tenn., Iowa, and Ariz.	90	100	0	–	0	–	0	–	90

Note: 1 cd/m² = 0.29 ft·L.

Figure 5. Dark surround.

Figure 6. Slightly illuminated surround.

numerical value that is converted to conventional luminance terms. The nonlinear relation of the EV and conventional luminance values (candelas per square meter or footlamberts) requires conversion of the EVs to conventional values and then the averaging of the results. The light weight, portability, and small size of these instruments make field use of them completely satisfactory. Instruments 2 and 3 have an internal illumination of the EV scale that is desirable for readings at low luminance levels.

The inclusion of only identical light sources in the surround is essential for similar readings. On the average, the measurement of a series of signs along a single route will produce similar data, although there will be some inevitable differences between observers measuring the same sign because they may stop at different distances from it or at differing offsets with respect to the traveled way.

The most satisfactory method of measuring surround data is as follows:

1. The observations are made from a vehicle in the traveled lane, while the driver's normal viewing point is maintained. A large offset, as within a driveway or parking area, displaces the sign with respect to its normal surround and leads to a slightly different surround that may have more or less luminance than does the actual one.

2. The observations are made from distances of approximately 90 m (300 ft) for the smaller regulatory and warning signs on the shoulder and approximately 180 m

Figure 7. Moderately illuminated surround.

Figure 8. Brightly illuminated surround.

Table 5. Descriptions of sign surrounds illustrated in Figures 5, 6, 7, and 8.

Surround	Description	Avg Luminance range (cd/m²)	No. of Readings
Dark	Occasional street or highway lighting; few commercial signs or other light sources; generally dark behind sign	<2.0	112
Slightly illuminated	Some street lighting or highway luminaires; occasional commercial signs and other moderately intense light sources adjacent to and behind sign	2 to 6	21
Moderately illuminated	Continuous street or highway lighting; frequent commercial signs adjacent to and behind sign	6 to 17	67
Brightly illuminated	Bright commercial signs, luminaires, and other light sources immediately adjacent to and behind sign	>17	40

Note: 1 cd/m² = 0.29 ft·L.

(600 ft) for the larger guide signs. These correspond to the distances in which motorists must observe signs and still have sufficient time to read them.

3. The area of interest is that immediately around the sign, and four representative measurements, as illustrated in Figure 4, are desirable for averaging to obtain the surround luminance.

Instruments 1 and 3 showed good agreement in a field-comparison measurement of approximately 30 sign surrounds.

Surround measurements of 166 signs were made by using instrument 3. The areas in which these measurements were made include (a) dark, rural roads; (b) illuminated, depressed freeway sections; (c) illuminated, at-grade freeway sections in both rural and urban areas; (d) suburban shopping centers; (e) downtown local streets; and (f) older, built-up highways. These areas are typical of those that can be found anywhere in the United States. Readings taken in six states and in six cities are given in Table 4.

Typical sign surrounds are illustrated in Figures 5 through 8. Their descriptions and measurements are given in Table 5.

RESULTS AND CONCLUSIONS

The dark category has few lights, and these are not troublesome. Reflective signs generally have sufficient contrast against this light level for good visibility.

The slightly illuminated category is variegated and involves light sources that diminish sign performance. At the lower end of the range, there may be one or more street-lighting luminaires close to a sign, but the other side of the sign will have good contrast against a dark background. At the upper level of the range, there will be street lighting, traffic signals, and distant commercial signs or displays. These additional light sources diminish the attention-catching value of the sign.

The moderately illuminated category is consistently troublesome above the 6.0-cd/m^2 (2.0-ft•L) level, where the detection of traditional traffic-control signs becomes difficult. The contrast is frequently negative; i.e., the sign is darker than the light sources around its edge.

The brightly illuminated category presents a highly variegated background that consists of street lighting, large areas of internally illuminated commercial signs, frequent intense sources such as spotlights and large incandescent bulbs, and static and flashing displays, all close to each other and to the road edge.

The evaluation of nighttime sign surrounds by measurements made with a spot-reading photographic light meter is suggested. These instruments have a relatively close correlation with laboratory-quality instruments and measure a 1° area. This corresponds closely to the critical area at the center of the visual field where maximum visual acuity is most seriously affected by proximate sources of glare, which reduce the legibility distance and require higher luminance of the sign.

The evaluation of numerous surrounds showed that terms such as dark, rural and bright, downtown, although illustrative, are misleading. There are many rural locations where distant sources of glare make the area as luminous as heavily developed, bright, downtown areas. Similarly, bright, downtown areas have frequent dark sections that are equivalent to dark, rural areas. The lack of correlation with traditional verbal descriptions is common.

There are also many locations where one side of a tangent section of a roadway has frequent glare sources, but the opposite side is relatively dark, e.g., the opposite sides of a roadway approaching a commercial development.

In many cases, overhead signs are seen against the night sky, which is a dark surround, but shoulder-mounted signs on the same road may have a moderate or bright surround. The pattern of night lighting is frequently concentrated along road edges and provides little above the road. This requires separate evaluation of overhead and shoulder-mounted signs. Separate evaluation is also necessary on curved roadways. Both moderate and abrupt changes in horizontal or vertical alignment may also align traffic signs with glare sources in an otherwise dark environment, and individual assessment of such situations is required to determine the exact location and extent of the surround luminance.

Future research should develop recommendations for appropriate sign-luminance levels for various night surrounds so that the highway engineer can design signs that are appropriate to their night surrounds.

REFERENCES

1. J. S. Smyth. The Brightness and Legibility at Night of Road Traffic Signs. Trans., Illuminating Engineering Society, Vol. 12, No. 4, 1947.
2. R. J. Lythgoe. The Measurement of Visual Acuity. Medical Research Council, Special Rept. Series 173, 1952.
3. T. M. Allen and A. L. Straub. Sign Brightness and Legibility. HRB, Bulletin 127, 1955, pp. 1-14.
4. T. M. Allen, F. N. Dyer, G. M. Smith, and M. H. Janson. Luminance Requirements for Illuminated Signs. HRB, Highway Research Record 179, 1967, pp. 16-37.
5. T. W. Forbes. Luminance and Contrast for Sign Legibility and Color-Recognition Results and Applications. Michigan State Univ., East Lansing, June 1976.
6. Deutsche Normen, DIN 67521, 1967.
7. The Construction of Road Traffic Signs and Internally Illuminated Bollards. British Standards Institution, BS 873, Pt. 1, 1970.
8. Standard Specifications for Construction of Roads and Bridges on Federal Highway Projects. U.S. Department of Transportation, FP-74, 1974.
9. Sheeting and Tape, Reflective: Nonexposed Lens, Adhesive Backing. Federal Specification L-S-300B, July 1974.
10. D. R. Hanson and H. L. Woltman. Sign Backgrounds and Angular Position. HRB, Highway Research Record 170, 1967, pp. 82-96.
11. W. P. Youngblood and H. L. Woltman. A Brightness Inventory of Contemporary Signing Materials for Guide Signs. HRB, Highway Research Record 377, 1971, pp. 69-91.
12. M. Luckiesh. Brightness Engineering. Illuminating Engineering, Feb. 1944.
13. T. M. Matson, W. S. Smith, and F. W. Hurd. Traffic Engineering. McGraw-Hill, New York, 1955.
14. B. D. Greenshields. The Driver. In Traffic Engineering Handbook (J. E. Baerwald, ed.), ITE, 1965.

Publication of this paper sponsored by Committee on Visibility.

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

Using Encapsulated-Lens Reflective Sheeting on Overhead Highway Signs

R. N. Robertson and J. D. Shelor, Virginia Highway and Transportation Research Council, Charlottesville

This paper summarizes a study on the use of encapsulated-lens reflective sheeting on overhead signs without external illumination. The existing signs on the Interstate and many primary highways in Virginia were inventoried to determine the percentage of them that would meet the criteria for visibility-recognition distance so that their illumination could be eliminated if they were refurbished with encapsulated-lens sheeting. The plans for several proposed sign-lighting projects were also reviewed for the same criteria. Data relative to the installation, energy, and maintenance costs for lighting overhead signs were also collected. It was concluded that illumination could be eliminated on approximately 45 percent of the existing signs and 50 percent of the proposed ones. The anticipated benefits include monetary and energy savings, reduction in the exposure of maintenance personnel to hazardous working conditions, and improved services to motorists.

The brightness of encapsulated-lens (high-intensity) reflective sheeting is superior to that of the enclosed-lens sheeting that is presently used on overhead traffic signs (1, 2, 3, 4). Consequently, the performance of this material was evaluated by the Virginia Department of Highways and Transportation (VDHT) to determine the feasibility of using it on overhead signs without illumination, and it was concluded that the use of encapsulated-lens sheeting would allow the elimination of the external lighting on many overhead signs without adversely affecting the service to the motoring public (5).

Subsequently, a joint study team from the Office of Engineering and Traffic Operations and Research and Development of the Federal Highway Administration (FHWA) evaluated the performance of encapsulated-lens sheeting in five states (6), and the FHWA removed the use of encapsulated-lens sheeting from the experimental category and established guidelines for the elimination of external lighting on overhead guide signs that are made with encapsulated-lens material.

The use of encapsulated-lens sheeting and the elimination of lighting on many overhead signs should be advantageous to many transportation agencies. Intuitively, the benefits would appear to include monetary savings, energy conservation, increased safety for maintenance personnel, and improved service to motorists. However, the consideration of these probable benefits generates questions such as the following: What percentage of the signs in Virginia meet the criteria for the elimination of lighting? What is the installation cost for lighting? What is the energy cost for lighting an overhead sign? What is the maintenance cost for the lighting on a typical overhead sign?

The purpose of this study was to answer the questions above; it was not intended to provide an economic analysis. The main objectives were to

1. Determine the percentage of existing and proposed overhead signs that meet the criteria for the elimination of lighting by the use of encapsulated-lens materials,
2. Obtain cost estimates for the installation of lighting on a typical overhead sign,
3. Obtain cost estimates for the energy used in illuminating overhead signs, and
4. Obtain cost data for the maintenance of the lighting fixtures on overhead signs.

Because of personnel and time constraints, the study was restricted to the Interstate and primary highway systems. Random samples of statewide data were collected, but it was impossible to obtain complete data on all the overhead signs in the state.

METHODOLOGY AND RESULTS

The first phase of the study was divided into four major tasks.

Sign Survey

One of the criteria established by the earlier study (5) is that the illumination can be eliminated from an encapsulated-lens sign on a freeway that has a straight approach equal to or greater than the visibility-recognition distance. The use of a model developed by Forbes (7) showed that the calculated visibility distance for the overhead signs on a freeway is about 335 to 366 m (1100 to 1200 ft). In terms of time, this allows a motorist traveling at freeway speeds 13.5 s to observe a sign after detecting it. On the assumption that this amount of time is sufficient for the motorist to identify and read the sign, the relationship of speed and visibility distance shown in Figure 1 was developed and used on roadways that had speed limits lower than those on freeways.

All of the overhead signs on the Interstate system, but only a sample of those on the primary and secondary roadways and city streets, were surveyed. The following data were recorded: (a) location of sign structure, (b) number of signs per structure, (c) type and number of lighting fixtures, (d) straight approach distance, (e) type of roadway, and (f) posted speed limit. The existing signs on roadways that are under construction were not inventoried, but the signing plans of several proposed projects were reviewed to estimate the percentage of signs that could be built with encapsulated-lens sheeting and without illumination.

The inventory of existing overhead signs on the Interstate highways given below shows that there are 271 sign structures on which 576 signs are placed.

		Other Roadway	
Item	Interstate	Counted	Estimated
Structures	271	199	265
Curved approaches	149	110	146
Straight approaches	122	87	116
Signs	576	446	594
Signs per structure	2.13	2.24	

Of these structures, 122 (45 percent) are located on straight roadways and meet the visibility-recognition criterion for the elimination of lighting by the use of encapsulated-lens reflective sheeting.

Approximately 75 percent of the total sign structures on other streets and highways were also surveyed. As shown above, 87 (43.8 percent) of the 199 structures were located on straight approaches.

Since the luminances of signs located on straight roadways are greater than those of signs located on curved roadways, in recent years designers have placed overhead signs on straight approaches whenever possible. The inventory given below of proposed sign structures on four construction projects shows that 50 percent of them are on straight approaches.

Item	Interstate (I-495)	Primary Roadway
Structures	148	10
Curved approaches	74	5
Straight approaches	74	5
Signs	231	21
Signs per structure	1.56	2.1
Light fixtures	580	62
Fixtures per sign	2.51	2.95

The inventory given below of lighting fixtures on existing roadways shows the number and variety required (1 m = 3.3 ft).

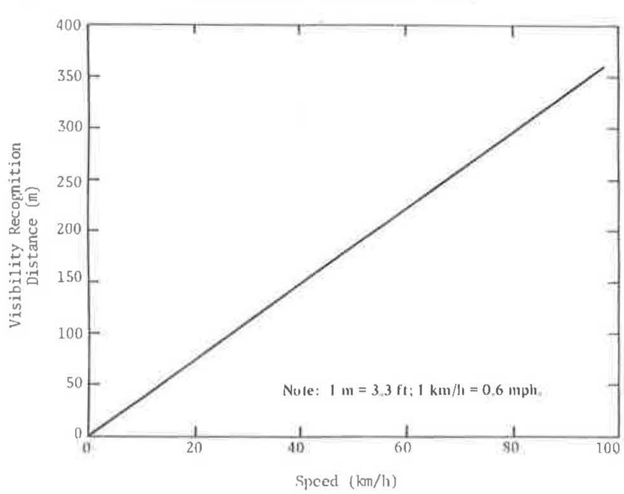
Type of Fixture	Interstate	Other Roadway	
		Counted	Estimated
1.22 m fluorescent	1027	834	1112
1.83 m fluorescent	228	230	306
2.44 m fluorescent	23	3	4
Mercury-vapor	481	240	320

The number of structures supporting these fixtures and the number of signs lighted by them are given below.

Type of Fixture	Roadway	Structure		Signs	
		Counted	Estimated	Counted	Estimated
Fluorescent	Interstate	158	–	392	–
	Other	142	189	337	449
Mercury-vapor	Interstate	96	–	150	–
	Other	44	59	82	109
None	Interstate	17	–	34	–
	Other	13	17	27	36

There are about 800 mercury-vapor fixtures and 2700 fluorescent fixtures [totaling 3658 m (12 000 ft)] in service in Virginia. To illuminate the average overhead sign requires 4.35 m (14.3 ft) of fluorescent lighting fixtures or 3.1 mercury-vapor fixtures. The majority of the signs are equipped with fluorescent lighting, but the newer installations include mercury-vapor fixtures because of their better performance characteristics.

Figure 1. Visibility-recognition distance versus speed.



Installation Cost

The majority of the overhead signs in Virginia are installed by an outside agency, and, unfortunately for the purpose of this survey, the payment for the entire structure is made on a lump-sum basis. Obtaining cost estimates for the installation of the lighting only on a typical overhead sign required contacting many sign contractors and consulting engineers and the Traffic and Safety Division of VDHT. The contractors were reluctant to discuss unit prices for lighting fixtures because of the fluctuations among projects and the dates of the work, the increasing costs of materials, and the fact that a small project has a high unit cost but a large project has lower unit costs. However, the contractors did indicate that the VDHT estimate of $400/fixture was conservative.

If, as shown in the inventory of proposed overhead signs, the average number of fixtures on a proposed sign-installation project is 2.55 and an average of 1.59 signs are planned for each structure, the average cost of lighting each structure will be $1600.

Overhead sign structures of the 2-pole span type cost an estimated $738/m ($225/ft) of span, and those of the cantilever type cost $800/m ($250/ft). One-third to one-half of these costs are for the walkways on which the light fixtures are mounted, and since few of the sign-maintenance operations are performed from the walkways, this additional expense is mainly for the mounting and maintenance of the lighting fixtures. Of the 148 structures proposed for I-495, 70 are of the cantilever type, 52 are span structures, and 26 are mou[illegible] bridges. The average lengths of the cantilever, span, and bridge structures are 8.69, 33.2, and 7.01 m (28.5, 109, and 23 ft) respectively. The average cantilever structure will cost $7125, and the average span structure will cost over $24 500. (The cost figures for the bridge-mounted signs were not available because these signs require special supports, but their costs are expected to be in the same range as those for the cantilever structures.)

Since 50 percent of the proposed structures on I-495 will be on straight approaches on which encapsulated-lens sheeting without illumination would provide adequate luminances, the elimination of the lighting fixtures would save more than $402 000 for the structures alone, and there would be an added saving of $118 000 for the lighting fixtures themselves. The net savings would be approximately $520 000 on this highway facility or an average of $7030/structure. These figures are conservative, because they were derived from cost estimates for a project that will require a large number of signs and is in an urban area where electrical service is readily available. The costs of illuminating signs in rural areas increase rapidly because of the long distances to power sources: In remote areas, service cable costs $6.56/m ($2.00/ft). Finally, there is an additional saving from the elimination of glare shields, which are not required on encapsulated-lens reflective-sheeting signs.

Energy Cost

Data on the cost of energy for lighting signs in various sections of the state were gathered and analyzed. The data include the annual electrical costs, the suppliers, the locations of structures, the number of signs per structure, and the type and number of lighting fixtures.

The costs of electrical energy varied widely throughout the state. In a few areas, the state government had a special rate that usually (in 1974) was less than 2 cents/kW. Some typical 1974 energy costs are shown in Table 1. At that time, the annual costs for fluorescent lighting on a typical overhead sign in Virginia were between $35.82

Table 1. Energy costs for overhead sign illumination.

Location of Structure	Signs on Structure	Lighting Fixtures on Structure	Length of Lighting Fixtures (m)	Annual Cost of Electricity ($)	Annual Cost of Electricity per Meter of Fixture ($)
I-64W at Parham Road, Henrico County	3	11	13.4	132.00	9.85
US-29N at US-15, Culpeper County	3	10	12.8	334.62	26.14
I-95S at Va-619, Prince William County	3	7	9.1	75.40	8.29
I-81S at Va-614, Botetourt County	3	6	7.9	75.20	9.52
I-64W at I-81, Augusta County	2	12	14.6	380.20	26.04
US-29S at Va-739, Amherst County	2	6	7.9	93.60	11.85
US-29N at US-60, Amherst County	2	7	9.1	96.00	10.55
I-81S at Va-381, Bristol	2	4	7.3	162.87	22.31
Total	20	63	82.3	1349.89	—
Avg	2.50	7.88	10.29	168.74	16.40

Note: 1 m = 3 ft.

Table 2. Maintenance costs for overhead sign illumination.

District	Signs	Person-Hours	Equipment and Labor Costs ($)	Material Costs ($)	Total Costs ($)	Unit Cost per Sign ($)	Remarks
Culpeper	—	3846	21 300	2900	24 200	—	State forces; includes traffic control
Salem	49	—	3 300	1050	4 350	89	State forces; includes traffic control
Richmond (I-64)	63	595.4	9 500	2600	12 100	192	Contract; excludes traffic control
Suffolk (I-44, 64, and 264)	142	475.0	14 100	4100	18 200	128	Contract; includes traffic control

and $113.02, with an average of $71.35, but these costs have increased greatly since then.

For example, the four 1.8-m (6-ft) fluorescent fixtures on the overhead signs located on I-81 at the Va-381 interchange in Bristol have been replaced by four mercury-vapor fixtures, for which the current electrical cost is $3.82/light/month and the anticipated annual cost is $183.36.

Maintenance Cost

Since the VDHT accounting system does not have a specific charge code for sign-lighting costs, the daily work records over a 12-month period in two highway districts were reviewed. The data recorded included the costs of labor, equipment, and materials for maintaining the sign lighting. Data including the number of signs maintained, the number of person-hours required, and labor, equipment, and material expenditures were also collected in two districts in which most of the maintenance operations were carried out by outside contractors.

A review of the daily work records relative to the maintenance cost for the illumination of overhead signs in the Culpeper and Salem Highway districts is given in Table 2. Unfortunately, the number of signs in the Culpeper District was not available, and unit maintenance costs per sign could not be calculated. However, there was an obviously large expenditure for the maintenance of sign lighting, and a three-person crew was assigned to this work. Approximately $25 000 and 3846 person-hours were expended, but these were not sufficient for an effective sign-illumination maintenance program.

In the Salem District, the maintenance of lighting on 49 signs costs approximately $4350 with a unit cost per sign of $89. However, because the majority of the overhead signs in the Salem District are located near the maintenance shops and therefore require little travel time and expense, these costs are considered to be minimal.

The maintenance work on many of the overhead sign lights in the Richmond and Suffolk districts is done by outside contractors, who are compensated for labor and equipment on an hourly basis and provide all traffic control during the maintenance operations. These contractors also bill VDHT for the materials used in the repairs of the sign illumination, and the cost of replacement parts supplied by them is approximately twice the cost usually paid by VDHT for identical items.

The maintenance of the lighting on 63 overhead signs on I-64 in the Richmond District costs more than $12 000, excluding traffic control. The unit cost per sign was $192. In the Suffolk District, maintenance of the lighting on 142 signs on I-44, 64, and 264 in the Norfolk area costs $18 000. The unit cost per sign, including traffic control, was $128. The other sign lights in these districts were maintained by state forces.

CONCLUSIONS

Because of the limitations of the data, especially those pertaining to the costs of installation, energy, and the maintenance of overhead sign lights, definitive conclusions as to the impacts of the elimination of lighting on encapsulated-lens signs cannot be made. Since the beginning of the energy crisis in the winter of 1973 and the addition of the fuel-adjustment charge, electrical rates have increased so rapidly that the establishment of a true indicator of the energy costs for a typical overhead sign is impossible. The maintenance-cost data were compiled from daily work records and do not necessarily reflect the total cost for maintaining the sign lighting. Frequently, additional crews are required for operations such as traffic control and the replacement of underground cable, and the costs of these activities may not be included in the data presented in this paper. The installation-cost data are also only estimates because contract prices were not available (and the contractors indicated that the estimates were low). Consequently, it is assumed that the foregoing analysis and the following general conclusions are conservative.

The sign survey showed that approximately 45 percent of the existing 1170 overhead signs are located on roadways that have straight approaches and thus that the lighting could be eliminated by the use of encapsulated-lens sheeting.

The annual cost of electricity for and maintenance of the illumination on the typical overhead sign varied between $124.82 for a sign maintained by state forces in an area with low electric rates to $305.02 for a sign maintained by a contractor in an area with high electrical rates. The average annual cost was $160.35/sign. This annual expense is greater than the additional investment required to build signs with encapsulated-lens reflective materials rather than with the conventional enclosed-lens sheeting. Because the service life of encapsulated-lens materials exceeds 10 years, a benefit-cost ratio greater than 10 to 1 can be anticipated for signs mounted on existing structures and refurbished with encapsulated-lens reflective sheeting.

If the existing 520 signs located on straight approaches were refurbished with encapsulated-lens materials and the lights disconnected, there would be an annual saving of approximately $83 000 in electrical and maintenance costs. This saving does not include other benefits, such as the reduced exposure of maintenance personnel to traffic, improved services to motorists, the availability of maintenance crews and equipment for other work, and the reduction in time required for night inspections to locate malfunctioning lights.

Eliminating the lighting on new overhead sign structures would result in enormous savings in installation costs. Because overhead signs are usually located on straight sections of roadways, the number of proposed signs that meet the visibility-recognition criterion is increasing. Fifty percent of these signs will be located on straight approaches, where the illumination could be eliminated if they were made with encapsulated-lens sheeting. On the sign project proposed for I-495, the elimination of lights on the overhead structures could save $7030/structure (less $400 to $500 for the additional expense of the encapsulated-lens sheeting). The saving for the entire project would be more than $500 000, and greater savings per structure could be anticipated on projects that require a small number of signs and in areas where the power sources are long distances from the overhead signs.

ACKNOWLEDGMENTS

The assistance given by the Traffic and Safety Division and by the construction districts of the Virginia Department of Highways and Transportation is gratefully acknowledged. This study was conducted under the general supervision of J. H. Dillard of the Virginia Highway and Transportation Research Council and was financed through the Highway Planning and Research Program of the Federal Highway Administration.

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

REFERENCES

1. Evaluation of Reflective Sign Materials. Research and Development Section, Louisiana Department of Highways, Jan. 1973.
2. R. L. Rizenbergs. High-Intensity Reflective Materials for Signs. Kentucky Department of Highways, Res. Rept., May 1973.
3. W. P. Youngblood and H. L. Woltman. A Brightness Inventory of Contemporary Signing Materials for Guide Signs. HRB, Highway Research Record 377, 1971, pp. 69-88.
4. R. N. Robertson. Use of High-Intensity Reflective Materials in Highway Signing, A Literature Review. Virginia Highway and Transportation Research Council, Aug. 1973.
5. R. N. Robertson. Evaluation of High-Intensity Sheeting for Overhead Highway Signs. Virginia Highway and Transportation Research Council, Dec. 1974.
6. Encapsulated-Lens (High-Intensity) Reflective-Sheeting Sign Material. Federal Highway Administration, Notice N5040.17, June 15, 1976.
7. T. W. Forbes. Factors in Highway Sign Visibility. Traffic Engineering, Vol. 39, No. 12, Sept. 1969.

Publication of this paper sponsored by Committee on Visibility.

Poor Visibility Under Low-Beam Headlights: A Common Cause of Wrong-Way Driving

N. K. Vaswani, Virginia Highway and Transportation Research Council, Charlottesville

Through selected case studies, this paper illustrates the way in which the inadequate visibility of road signs and pavement markings at night contributes to wrong-way driving. A concept termed the keg of legibility, which delineates the limits of nighttime visibility under low-beam headlights, is described. The application of the keg-of-legibility concept to the placement of signs, markings, and additional devices that help guide the motorist through the intersection of a four-lane divided highway and another road is discussed. Examples of wrong-way entry on roads having poor geometrics are used to emphasize the need for such guidance.

Surveys of wrong-way driving in Virginia since 1970 have shown that most of the wrong-way incidents originated at interchanges and intersections. A driver must be very carefully guided onto the correct ramp at an interchange or around the nose of the median when he or she is making a left turn at an intersection on a divided highway. Many information devices, such as signs and pavement markings, and other features such as curbs, often made conspicuous by color, are used to provide this guidance, but they are often not of maximum effective-

Figure 1. Ten-degree cone of vision.

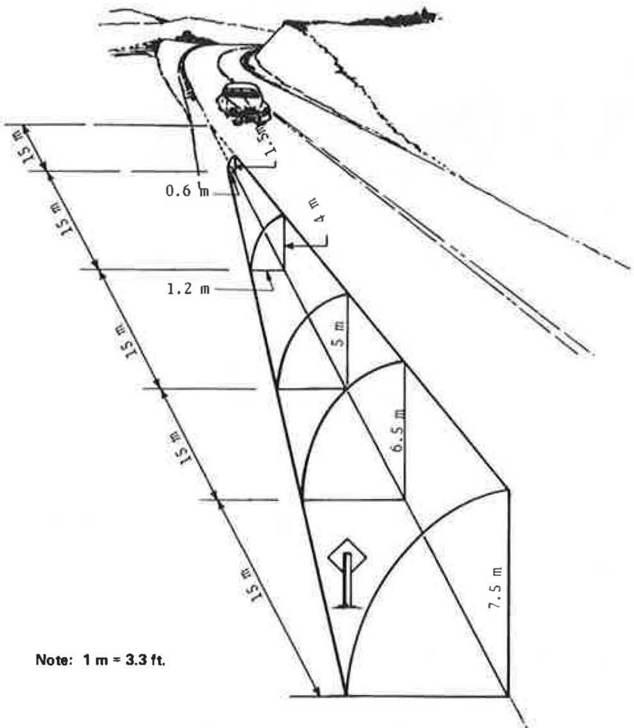


Figure 2. Keg of legibility of sign under low-beam headlights.

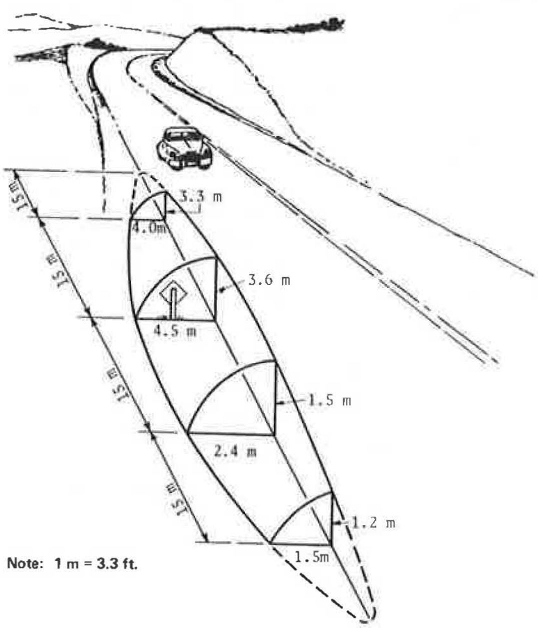


Figure 3. Daytime photograph of intersection of Interstate exit ramp and secondary road.

ness for a driver using low-beam headlights at night, either because of improper location of them or because of poor geometrics of the intersection.

NEED FOR DESIGN BASED ON NIGHT VISIBILITY

The longitudinal and transverse distances at which most traffic-control and guidance elements at intersections can be seen during the day are limited only by sight distances. However, at an unlighted intersection at night, these distances are limited by the following additional factors:

1. The area illuminated by the headlights, which are usually on low beam,
2. The size and shape of the object to be viewed,
3. The luminance and contrast of the roadway elements, and
4. The luminance of the details of the object and their contrast with their background.

Thus, roadway devices located on the basis of their daytime visibility and legibility may not be visible or legible at night, and the motorist is then left with only a limited number of them, or none at all, for guidance. Therefore, for full effectiveness, highway devices provided for guidance should be designed and located on the basis of their legibility under low-beam headlights at night.

The present concept of the field of vision of a driver is based on the cone of vision. According to Pignataro (1), the limit of far clear sight is that within a cone of 10 to 12°. Figure 1 shows a 10° cone of vision and the vertical and horizontal distances from the pavement edge within which, according to the cone concept, a sign would

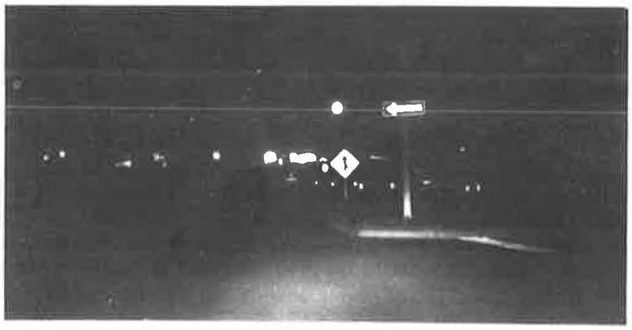
Figure 4. Nighttime photograph of intersection shown in Figure 3.

Figure 5. Keg of legibility for intersection shown in Figures 3 and 4.

Note: 1 m = 3.3 ft.

Figure 6. Daytime photograph of partial cloverleaf intersection of Interstate exit ramp and primary highway.

Figure 7. Nighttime photograph (low-beam headlights) of intersection shown in Figure 6.



be visible. This concept, however, is based on daytime vision.

The nighttime and daytime legibility of a 0.6 by 0.6-m (2 by 2-ft) reflectorized diagrammatic sign made of engineering-grade sheeting was evaluated. The sign was placed 0, 1.5, 3.0, or 4.5 m (0, 5, 10, or 15 ft) from the pavement edge, with its center 1.5, 2.4, or 3.3 m (5, 8, or 11 ft) above the road level. Nighttime and daytime photographs of it were taken at each combination of placements from distances of 15, 30, 45, 60, and 75 m (50, 100, 150, 200, and 250 ft). The lens of the camera was 1.2 m (4 ft) above the road surface and 2.7 m (9 ft) from the pavement edge. At night, low-beam headlights were used.

These photographs were projected in a darkened room before five persons who graded the legibility (poor, fair, good, or excellent) of the sign. The limits of good legibility in terms of depth, height, and distance from the pavement edge determined in this way are shown diagrammatically in Figure 2. This diagram shows that the zone of good legibility at night is not conical (as shown in Figure 1) but keg shaped. For example, a sign in a quadrant of an 4.5 by 3.6-m (15 by 12-ft) oval with its axis on the pavement edge should be legible to a driver 30 m (100 ft) away, and its legibility would still be good even if its distance from the pavement edge were increased to 1.5 m (5 ft). The maximum distances from the pavement edge within which the sign would still be legible to a driver 15 or 30 m (50 or 100 ft) away are 3 and 3.6 m (10 and 12 ft) respectively. Hence, for intersections at which the distance between the stopping point of the driver and the median (where the signs are located) is less than 30 m (100 ft), the maximum distance from the pavement edge for the placement of a sign can be taken as 3 m (10 ft).

This keg of legibility is that for a normal person driving with low-beam headlights in good weather conditions on a straight road. Its size will decrease with defective headlights, increased humidity, and fog and rain. It could, however, be increased by the use of high-intensity sheeting rather than engineering-grade sheeting.

ROADS INTERSECTING AT SAME ELEVATION

The following two cases, in which the drivers entered the exit ramp of the Interstate rather than the entry ramp, resulted from poor visibility of the signs and roadway markings.

Case 1—Intersection of Interstate Highway Exit Ramp and Secondary Road

Figures 3 and 4 are daytime and nighttime photographs of an exit ramp at the intersection of an Interstate highway and a secondary road, where a wrong-way entry occurred. Two things are evident from the photographs.

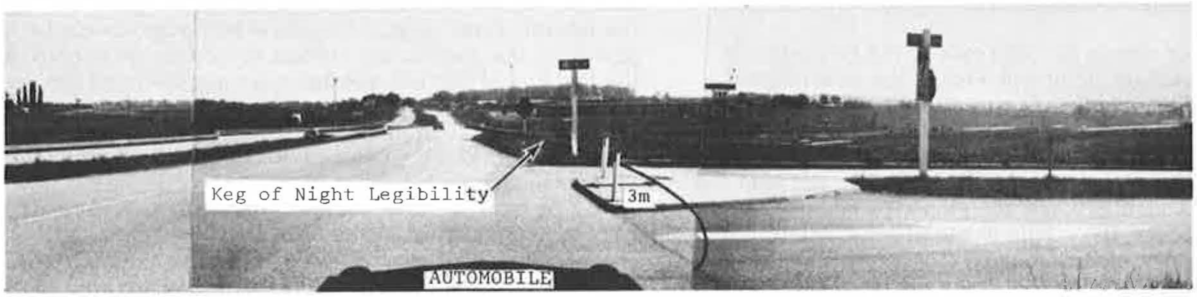
Figure 4 shows that, because of the restricted depth and width of vision at night, a driver with low external stimuli may be guided by the line at the edge of the pavement, which flares into the right lane. The continuation of this line straight across the ramp pavement might discourage a wrong-way entry at night. Or, an alternative way to prevent such an entry would be to place the stop line sufficiently close to the crossroad to put it within the zone illuminated by low-beam headlights, i.e., within the keg of nighttime legibility (2, 3). Either of these alternatives might channelize the movements of drivers, especially those with low external stimuli, and provide a pseudo-pavement-edge effect.

A comparison of Figures 3 and 4 shows that the one-way arrow sign and the stop line, which are visible to the daytime driver, are not visible at night. If drivers can function at night without the benefit of a particular sign, this sign evidently has no use during the daytime also. Hence, the locations of signs should be based more on their nighttime than on their daytime visibility. This one-way sign and the stop line should have been located within the keg of nighttime legibility, which is shown in Figure 5.

Case 2—Intersection of Interstate Highway Ramps and Primary Highway

Figure 6 is a daytime photograph of a partial cloverleaf (parclo) interchange between the exit and entry ramps of an Interstate highway and a divided primary highway, where a wrong-way entry occurred. The nose of the median between the exit and entry ramps is set back from the junction and, as shown in Figure 7, is not visible at night. If the nose were made visible at night it would show the separation between the exit and the entry ramps and thus would reduce the probability of a driver entering the exit rather than the entry ramp. The following improvements are recommended for a parclo inter-

Figure 8. Keg of legibility for intersection shown in Figures 6 and 7.



Note: 1 m = 3.3 ft.

Figure 9. Intersection of primary divided highway and secondary road at differing elevations (cross section across four-lane divided intersection extended into crossroad).

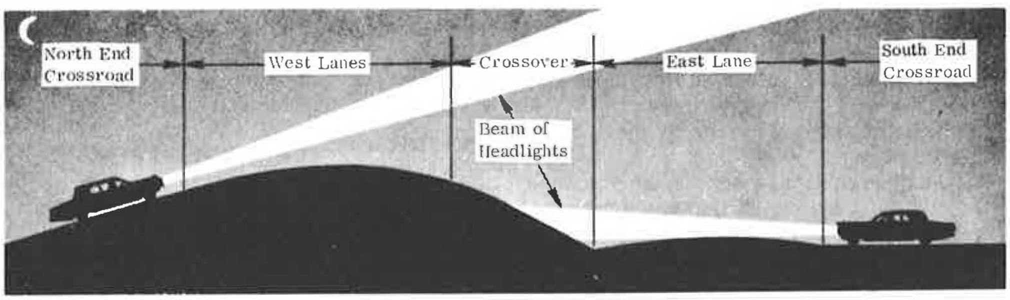
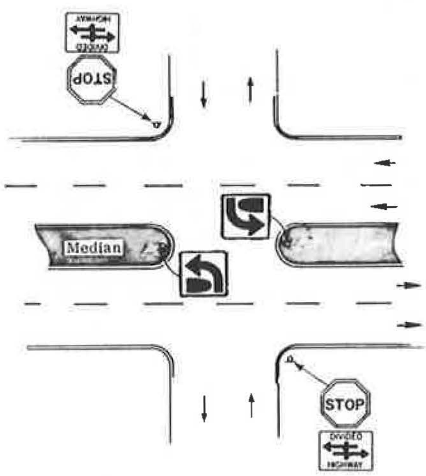


Figure 10. Recommended traffic signs for discouraging wrong-way entry.



change in which the exit and entry ramps are very close to each other.

1. The nose of the median should be extended to the edge of the crossroad so that it is within the keg of nighttime legibility, and it should be made of concrete and painted with reflective paint. It should also be marked by delineators, which should be within the keg of nighttime legibility. Figure 8 shows the suggested improvement with the portion of the nose that would be visible at night. Such a nose would provide proper visibility, separate the exit and the entry ramps, and fully channelize the exit ramp and thus discourage drivers from entering the exit ramp from the crossroad.

2. Either the pavement edge line should be continuous across the exit ramp, or the stop line should be close to the edge of the crossroad so that it is within the keg of nighttime legibility.

3. The pavement edge line should be flared into the entry ramp to encourage drivers to maneuver properly.

INTERSECTING ROADS AT DIFFERENT ELEVATIONS

General Cases

The two commonest problems involving the geometrics at nonlevel intersections are described below.

1. The crossroad slopes down from the divided highway. Sometimes the slope is so steep that little or no light from the headlights of an automobile approaching the divided highway falls on it. An example is shown in Figure 9.

2. The opposing lanes of the divided highway are at different elevations. A driver coming from the crossroad cannot see both sides of the divided highway with low-beam headlights and may consider it to be a two-lane road and the median to be the opposite edge of the road.

These problems are compounded when they are combined at one intersection. The steeper the downward slopes of the crossroads or the greater the difference between the elevations of the two opposite lanes of the divided highway, the poorer is the visibility.

Case 3—Intersection of Divided Primary Highway and Secondary Road

The intersection shown by the cross-sectional sketch in Figure 9 is the site of two wrong-way entries (both by nondrunken drivers): One entry was during the day from the northern side of the crossroad, and the other was at night from the southern side of the crossroad. As is shown in this figure, the northern side of the crossroad slopes down from the divided highway, and there is a considerable difference in elevation between the east and the westbound lanes of the divided highway. The southern side of the crossroad is, however, level with the eastbound lane of the divided highway.

Thus, a driver approaching the intersection from the northern side of the crossroad cannot see the highway. This kind of intersection could be improved as follows.

1. A driver should be informed of the geometry of the roadways before he or she enters the intersection. This can best be done by placing a diagrammatic sign depicting a divided highway intersection so that it is visible to a driver using low-beam headlights at night, when the need to know the geometry is greatest. The best location for this sign is below the stop sign and on the same pole (Figure 10). This would place it within 3 m (10 ft) of the edge of the lane and thus within the keg of nighttime visibility. Signs of this type have been installed on an experimental basis at intersections on 92 km (57 miles) of primary highways in Virginia. They have also been used in Delaware, where it is claimed that wrong-way entries have been reduced (according to a letter of August 7, 1976, from Raymond S. Pusey of the Delaware Bureau of Traffic to the Federal Highway Administration). It is not an international sign nor has it been approved for incorporation in the Manual on Uniform Traffic Control Devices by the U.S. Department of Transportation.

2. At intersections such as the one shown in Figure 9, the nose of the median that the driver must negotiate in making a left turn is not visible to him or her at night, and it may be necessary to provide guidance for this maneuver. This information is in addition to the divided-highway-intersection sign, and a suitable sign is also shown in Figure 10.

CONCLUSIONS

1. The locations of road signs and pavement markings should be designed on the basis of their nighttime visibility rather than their daytime visibility.

2. Diagrammatic signs should be used to provide guidance at intersections having poorly designed features, such as differences in elevation between the opposite lanes of four-lane divided highways, crossroads that slope down from divided highways, or wide crossovers that could lead to wrong-way entries. A diagrammatic sign depicting a divided-highway intersection should be placed below the stop sign at the junction of a crossroad and a divided highway to inform the driver of the geometry of the intersection. A diagrammatic turn sign should be placed on the nose of the median to inform the driver of the location of the left-median nose and the need for turning around it.

3. The application of the cone-of-vision concept for the placement of signs should be modified to include the keg-of-nighttime-legibility concept.

4. At intersections of crossroads and highway exit ramps, the marking on the edge of the pavement of the crossroad should be continued across the exit, or the stop line on the exit ramp should be brought within the keg of nighttime legibility of a driver on the crossroad.

5. On parclo interchanges having the exit and entry ramps very close together, the median should extend to the edge of the crossroad, and its nose should be made of concrete and painted with reflective material. This will make the nose conspicuous in the keg of legibility, show the separation between the exit and entry ramps near the crossroad, and channelize traffic from the exit ramp.

ACKNOWLEDGMENT

This study was conducted under the general supervision of J. H. Dillard of the Virginia Highway and Transportation Research Council, and was financed with state research funds. The opinions expressed are those of the author and are not necessarily those of the sponsoring agencies.

REFERENCES

1. L. J. Pignataro. Traffic Engineering, Theory and Practice. Prentice-Hall, New York, 1973.
2. N. K. Vaswani. Engineering Measures for Reducing Wrong-Way Driving. Virginia Highway and Transportation Research Council, Oct. 1975.
3. N. K. Vaswani. Case Studies of Wrong-Way Entries at Highway Interchanges in Virginia. TRB, Transportation Research Record 514, 1974, pp. 16-28.

Publication of this paper sponsored by Committee on Visibility.