ties for a particular site, a more reliable evaluation will result. In this case, however, the benefit-cost procedure of this paper provides a useful framework for evaluating the available alternatives.

The user should also recognize that some important considerations are beyond the scope of an economic analysis but may well have an important impact on the final decision. For example, the economic analysis does not completely reflect the role of operational flexibility in evaluating median treatments. An arterial highway with a two-way left-turn lane is far more flexible operationally than a highway with a median barrier. Such flexibility makes routine operation less restrictive since left-turns are not prohibited, and the treatment has better service capability under transient conditions such as roadway construction or a traffic accident. In this case, both the economic and operational considerations favor the same median treatment, but in other situations there may be trade-offs to be made by the decision maker. In short, the economic analysis is an extremely important part of the selection of an optimal median treatment, but other less quantifiable factors also deserve consideration.

ACKNOWLEDGMENTS

Much of the analysis in this paper was performed by Midwest Research Institute as part of a Federal Highway Administration contract. We acknowledge this sponsorship but also state that the findings and conclusions of this paper are our own and do not necessarily represent the views of the Federal Highway Administration.

REFERENCES

1. J. A. Azzeh, B. A. Thorson, J. J. Valenta, J. C. Glennon, and C. J. Wilton. Evaluation of Techniques for the Control of Direct Access to Arterial Highways. Federal Highway Administration, U.S. Department of Transportation, Rept. FHWA-RD-76-85, Aug. 1975.
2. J. C. Glennon, J. J. Valenta, B. A. Thorson, and J. A. Azzeh. Technical Guidelines for the Control of Direct Access to Arterial Highways—Volumes 1 and 2. Federal Highway Administration, U.S. Department of Transportation, Repts. FHWA-RD-76-86 and FHWA-RD-76-87, Aug. 1975.
3. Manual on Uniform Traffic Control Devices for Streets and Highways. Federal Highway Administration, U.S. Department of Transportation, 1971.
4. T. H. Mulinazzi, and H. L. Michael. Correlation of Design Characteristics and Operational Controls With Accident Rates on Urban Arterials. Joint Highway Research Project, Purdue Univ. and Indiana State Highway Commission, 1967.
5. J. W. McDonald. Relation Between Number of Accidents and Traffic Volume at Divided Highway Intersections. HRB, Bulletin 74, 1953.
6. G. M. Webb. The Relationship Between Accidents and Traffic Volumes at Signalized Intersections. Proc., ITE, 1955.
7. An Evaluation of a Safety Project. Traffic and Safety Division, Michigan Department of State Highways, Lansing, Dec. 1968.
8. Evaluation of Four Safety Projects. Traffic and Safety Division, Michigan Department of State Highways, Lansing, TSD-G-207-2, 1972.
9. J. C. Ray. Evaluation of a Two-Way Median Left-Turn Lane. Traffic Engineering, March 1961.
10. R. B. Sawhill and D. R. Neuzil. Accidents and Operational Characteristics on Arterial Streets With Two-Way Median Left-Turn Lanes. HRB, Highway Research Record 31, 1963, pp. 20-56.
11. R. D. Lewis and others. A Study of Two-Way Left-Turn Lanes by Technical Council Committee No. 10. ITE, Technical Notes, Dec. 1976.
12. M. R. Hoffman. Two-Way, Left-Turn Lanes Work! Traffic Engineering, Aug. 1974.
13. R. C. Thomas. Continuous Left-Turn Channelization and Accidents. Traffic Engineering, Dec. 1966.
14. Need to Improve Benefit-Cost Analyses in Setting Motor Vehicle Safety Standards. National Highway Traffic Safety Administration, U.S. Department of Transportation, July 1974.
15. T. C. Thomas and G. I. Thompson. The Value of Time Saved by Trip Purpose. Stanford Research Institute, 1970.

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

Highway Design Consistency and Systematic Design Related to Highway Safety

John C. Glennon, Transportation Consulting Engineer, Overland Park, Kansas
Douglas W. Harwood, Midwest Research Institute, Kansas City, Missouri

This paper proposes a more systematic approach to highway design for achieving consistent designs to meet the needs of drivers. It is intended as a catalyst toward promoting optimal improvements of existing facilities. Its nature is conceptual. The topics covered include (a) a critique of current practices; (b) the evolution of highway design; (c) objectifying the design process; (d) consistency of design in relation to driver expectancy; (e) application to achieve design consistency; and (f) developing a cost-effectiveness methodology.

For almost 4 decades, highway designers have relied on criteria presented in a series of design policies of the

American Association of State Highway and Transportation Officials (AASHTO). Although these publications provide a unique framework for geometric design, they neither treat geometric design as a systematic process nor provide any insights on designing highways to meet the critical needs of drivers. The AASHTO design policies have often led to inconsistently designed highways. Conceived as a way of communicating standards of good practice, these policies have often become the sole authorities. When asked about the adequacy of a design, some designers say "It's consistent with the AASHTO blue book" rather than "It meets the needs of the driver."

As the recent AASHTO "3R Guide" (1) shows, the basic scenario of the highway community is rapidly changing from a massive road-building campaign to a decided attempt to optimize the traffic safety and service of existing highways. Although many design errors are "poured in concrete," this changing emphasis provides an outstanding opportunity to improve existing highways so they are more consistent with the needs of the driver. But this goal can only be accomplished if the design process is objectified to the extent that it maximizes the effectiveness of design improvements subject to funding constraints.

EVOLUTION IN HIGHWAY DESIGN

Between the inventions of the wheel and the automobile, the primary concern of road builders was "getting the road user out of the mud." Only the structural aspects of design were considered. In the 1920s, when a personal automobile became a reality for many people, there began the evolution of a highway design technology of which many remnants remain. Most of the early highway design engineers came from railroad engineering backgrounds.

As highway transportation developed in the 1930s (aided particularly by government employment-support programs), more and more paved roads were built. By the late 1930s, the number and speeds of vehicles began to multiply. With these trends came frequent traffic jams and large increases in highway fatalities.

In 1937, as a reaction to these highway transportation problems, the American Association of State Highway Officials (AASHO) organized the Special Committee on Administrative Design Policies. The purpose of this committee was the formulation of administrative policies aimed at stimulating uniform practices of good highway design that would result in maximum safety and usefulness. Between 1938 and 1944, this committee formulated the following seven policy statements: A Policy on Highway Classification, September 16, 1938; A Policy on Highway Types (Geometric), February 13, 1940; A Policy on Sight Distance for Highways, February 17, 1940; A Policy on Criteria for Marking and Signing No-Passing Zones for Two- and Three-Lane Roads, February 17, 1940; A Policy on Intersections at Grade, October 7, 1940; A Policy on Rotary Intersections, September 26, 1941; A Policy on Grade Separations for Intersecting Highways, June 19, 1944; and A Policy on Design Standards—Interstate, Primary and Secondary Systems. Many of the criteria presented in these policies still undergird current AASHTO design policy manuals. These criteria, of course, were based on the vehicle performance, highway design, and traffic operations of the 1930s. As a result, the validity of their application in current highway design technology may be questionable.

As an example of the mismatch between design standards and current highway operations, consider the example of the design and operation of passing zones. First, the design of passing sight distance (2) only indirectly considers the design of usable passing zones. The second inconsistency is that the design for passing sight distance and the striping of highways for no-passing zones are based on entirely different criteria. The current MUTCD (3) standards for no-passing zones (which indirectly set the dimensions for passing zones) are based on criteria presented in the 1940 AASHO policy (4). Unlike the current design for passing sight distance, which uses a constant 16.1-km/h (10-mph) speed difference between passing and passed vehicles for all design speeds, the sight distance for striping is based on speed differentials that range from 16.1 km/h (10 mph) at a 43.3-km/h (30-mph) design speed to 40.2 km/h (25 mph) at a design speed of 112.7 km/h (70 mph). These criteria are considerably more liberal (and more hazardous) than the design criteria (5).

Not only is the validity of current design standards in question but, more important, geometric design problems are also compounded by the lack of a systematic approach to highway design. Present methods of design are often based on solutions to old problems rather than the specific nature of the problem at hand. In addition, because of the complexity of highway design, the design tasks are generally assigned to seemingly independent teams, which ignores the basic principles of system design optimization. Although direct lines of communication may exist between task teams, the lack of defined responsibility and authority toward the total system design may prevent a solution close to the optimum.

Recently, increasing emphasis has been given to the systems engineering approach to design. This is a creative form of problem solving that emphasizes the total design or task rather than merely considering the efficiency of each component part. The primary principle in applying the systems approach to design is to maximize system performance for a given cost or to minimize cost for a given performance. This general approach, of course, is not new. What is new is that the systems approach is completely rational rather than intuitive and uses such formalized techniques as game theory, queuing theory, linear programming, dynamic programming, control theory, critical path methods, network theory, and various optimization techniques.

In the past, the complexity of the highway systems design process often forced highway administrators to decompose the process unnaturally into noninteractive tasks, ignoring many of the necessary feedback aspects of the process. Unfortunately, the highway engineering community has not had the necessary tools to consider all of the interactions, let alone objectively weigh alternative designs, in coordinating the data and performing the design.

Now that the Interstate system is nearly complete and there is a trend toward improving the safety and usefulness of the 5 968 000 km (3 700 000 miles) of existing highway network, it is past time to develop an objective design process whereby the design engineer can both design new facilities and optimize future efforts to improve the performance of the existing system.

OBJECTIFYING THE DESIGN PROCESS

To avoid some of the highway design problems of the past requires a comprehensive description of the highway design process. In other words, the total process must be completely defined from setting goals to achieving the completed design (or redesign) of a highway. The entire, conceptualized design process is shown in Figure 1 and discussed below. An appreciation of the relations and interactions shown in Figure 1 is the first step toward making each element of the idealized design process concrete rather than abstract. A major research

Figure 1. The highway design process.

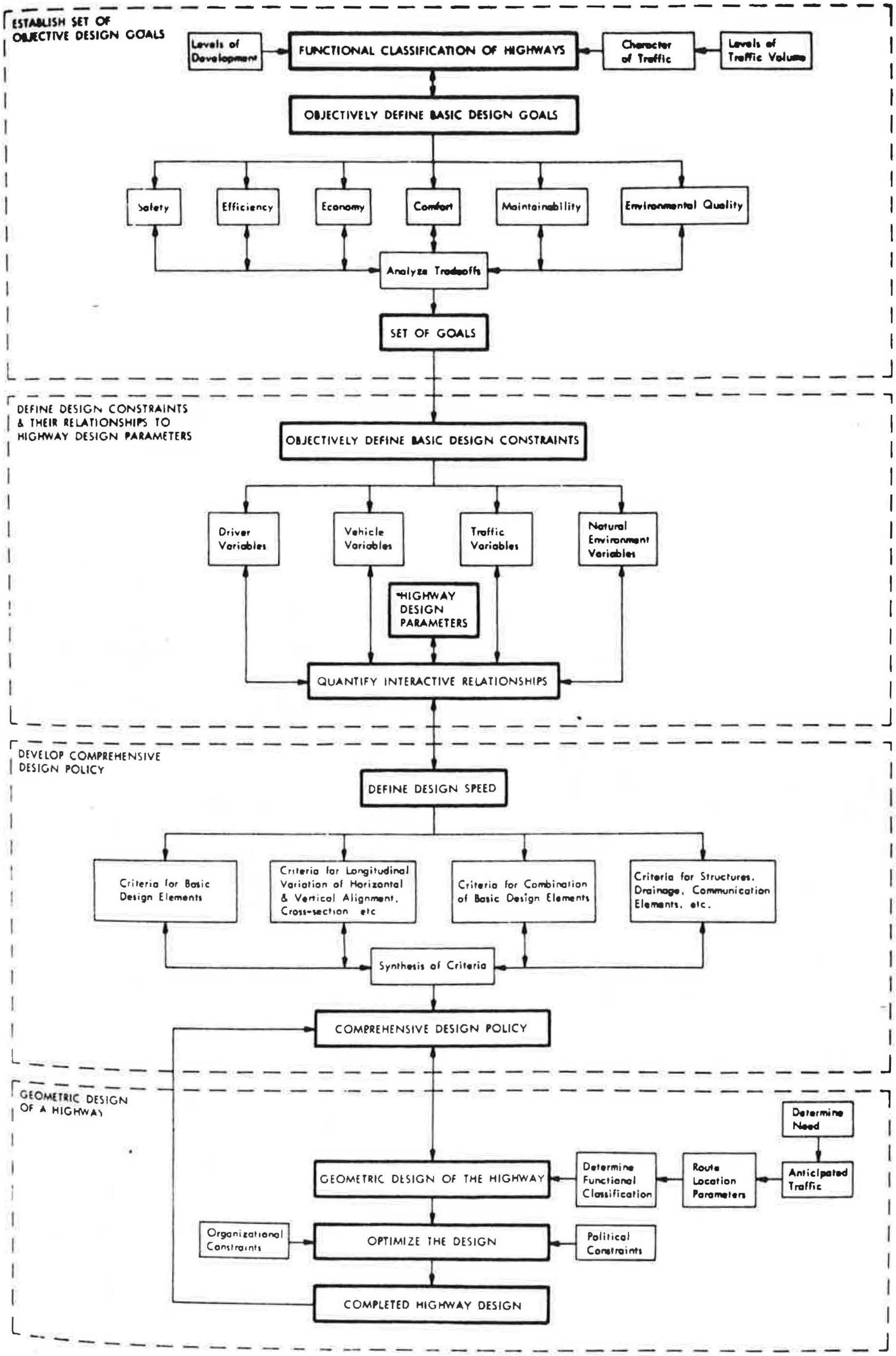
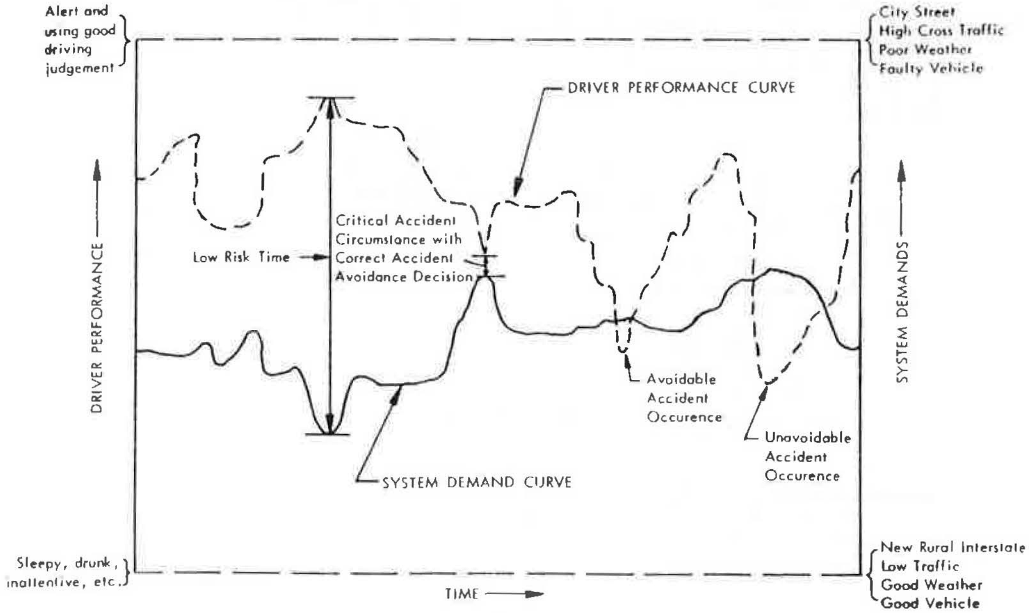


Figure 2. Conceptualized relation between driver performance and highway system demands in creation of accident circumstances.



effort is necessary to develop practical, yet valid, procedures and criteria for each element of the design process. The following discussion of the design process suggests an approach that addresses the need for specific design criteria, performance measures (measures of effectiveness), and decision-making tools.

The apex of the objective design process is the requirement that desired goals be defined and completely quantified. In addition, of course, these goals must be defined within the framework of a functional classification of highways. This points to a primary weakness of the AASHTO policies. Although they name the goals of safety, efficiency, economy, and comfort, they do not operationally define these goals.

The first part of objectifying the design process, therefore, requires a format for functional classification of highways and the formulation of a framework for operationally defining the goals of highway design in each functional class. The functional classification should consider the trade-offs between the functions of traffic service and land access, including rural or urban development needs and the level and type of traffic to be served.

The second major step in defining the design process shown in Figure 1 is an objective description of the basic constraints of driver, vehicle, traffic, and environmental characteristics and the interactive relations among these characteristics and between them and roadway characteristics. It is also important to identify how these constraints and their interactions set the requirements for the development of design criteria.

In developing design criteria that are functionally related to the design constraints, the real solution is one of matching the limited sensory and motor capabilities of the driver to the requirements of the driving task for various combinations of vehicle, roadway, traffic, and environmental constraints. Figure 2 shows conceptually that the performance of most drivers is usually adequate to the demands of the highway system. Accidents occur when either (a) driver performance falls below the level required by the system at that time or (b) system demands exceed driver performance at that time. In developing geometric design criteria, therefore, a basic principle should be to avoid peaks in the system demand curve created by inconsistency in design.

In the design process, a lack of understanding of basic design constraints and how they affect the solution contributes to piecemeal solutions that prevent optimization. The current approach tends to ignore the consistency of various combinations of design elements and thus oversimplify the process and limit the reliability of relations for most design purposes. But the primary reason for the lack of useful and definitive relations between design criteria and basic operational constraints is that these definitions depend on the complex interactions between the components of the highway transportation system, between their attributes, and between these and their environment. Until the significant interactions in the system can be quantified, reliable design criteria cannot be established.

The next major step in the design process (Figure 1) involves defining design speed as a function of design goals and constraints for each of the functional classes of highway. Without question, the "design equation" is most sensitive to vehicle speed—not only because the ability to stop or corner is a function of the square of speed but also because the impact forces of a collision are also a function of the square of speed.

AASHO design policies define design speed as "the maximum safe speed that can be maintained over a specific section of highway when conditions are so favorable that the design features of the highway govern" (2). This definition is abstract and does not lend itself to being an objective basis for design. It is difficult to imagine, under conditions "so favorable" and with modern design standards of 3.6-m (12-ft) lanes, flat cross slopes, and relatively flat grades, that any design feature other than horizontal curvature could govern maximum safe speed. Actually, in a physical sense, this is true. If driver, vehicle, traffic, and environmental constraints are eliminated from the design equation, the only design feature that physically governs maximum safe speed (for modern highway designs) is horizontal curvature. If this were true in an operational sense, the speed for long, level, tangent sections would be unrestricted and, where horizontal curvature was introduced, the concept of an overall design speed for that facility would be incongruous.

What is required is an operational definition of design speed that encompasses driver, vehicle, roadway, environmental, and traffic constraints and their relations to the design of an efficient, safe, and economical highway facility. To achieve this basis, for example, the designer requires knowledge of the characteristics of a "design vehicle" and how they relate to vehicle stability at various speeds—e.g., aerodynamics, suspension, weight, weight distribution, steer angle related to turning radius, accelerative capabilities, and braking capabilities.

The next major step is to define the design criteria objectively. The different kinds of criteria apply to the specification of the basic design elements, the longitudinal variation of horizontal, vertical, and cross-sectional elements, and the combinations of design elements (in general but also for special locations such as intersections, interchanges, and weaving sections). The process of developing design criteria involves analyzing the criticality of the interactive relations between the design constraints and the design elements for various highway speeds and selecting that level of criticality that limits the probability of an undesirable event (e.g., accident or congestion).

Synthesis is an important and necessary part of this development. Complete and comprehensive documentation of data is of little use unless it can be synthesized into a usable body of knowledge. By means of this kind of synthesis, sensitivity analysis can be performed to identify the more significant parameters that affect the safety effectiveness of any design improvement.

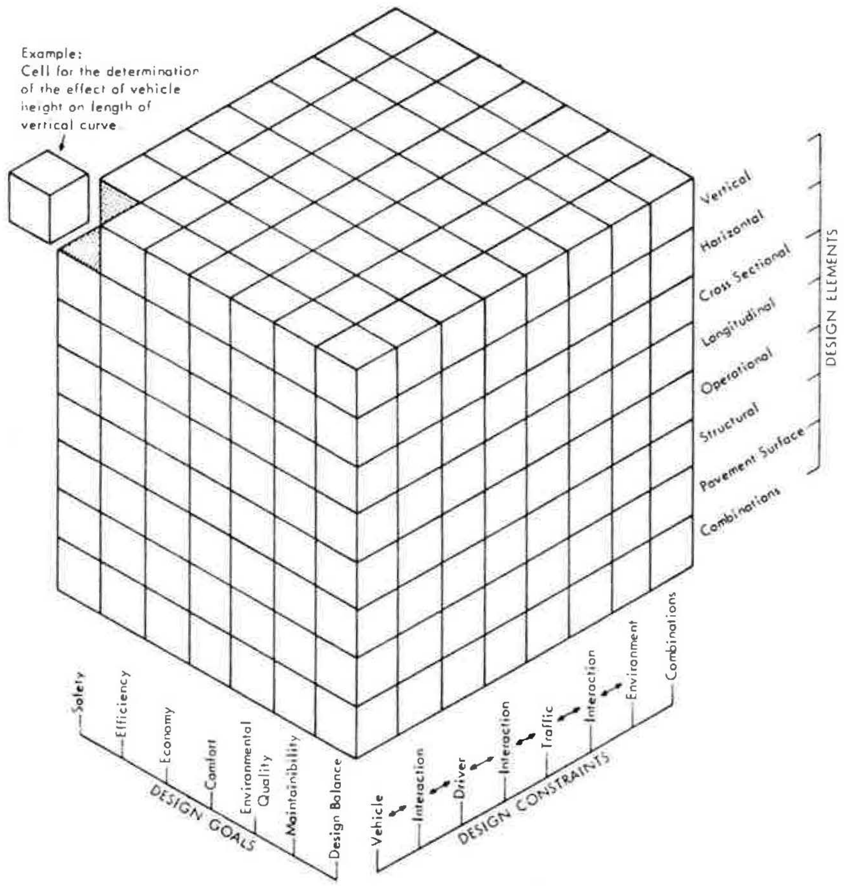
Figure 3 shows the general matrix of analysis. The necessary synthesis of data and information on interactive relations involves the following five basic steps:

1. Define a measure of hazard;
2. Formulate multidimensional data matrixes;
3. Statistically select an appropriate hazard value from data elements in each matrix cell;
4. Apply statistical procedures to predict expected hazard values for empty matrix cells; and
5. Reiterate the synthesis process, combining matrixes for higher orders of development.

First, we must define a measure of hazard so that the effect of varying dimensions of highway design elements, and combinations thereof, can be objectively evaluated (this step is described further below). Second, for each value of a design element, multidimensional data matrixes of hazard measures are classified by incremental values for the various combinations of the design constraints. The class range for each design element or design constraint in the matrix is then determined by analyzing the sensitivity of the dependent hazard measures to variations in the values of the design elements and design constraints. Third, within each cell of each matrix, the data elements (if there are more than one) are statistically analyzed to select the appropriate hazard value for that cell. Fourth, statistical procedures (analysis of variance, multiple regression, and so on) are applied to each data matrix to predict the expected hazard values for any empty matrix cells. And, finally, the synthesis process is reiterated, and successively higher orders of development are achieved by combining appropriate matrixes (submodels) into more inclusive matrixes.

A measure of hazard must be defined so that the effect of varying dimensions of highway design elements,

Figure 3. General matrix for synthesis of interactive relations.



and combinations thereof, can be evaluated by some criterion of "good." At any location, the degree of accident hazard is a function of two variables: accident frequency and accident severity. If two locations have the same accident frequency, the one that has the lower accident severity is less hazardous. If two locations have the same accident severity, the one that has the lower accident frequency is less hazardous. Thus, neither accident severity nor accident frequency can serve alone, but both must be integrated into one criterion.

The degree of accident hazard can be defined in several ways. It is a measure of the potential for a particular highway location to produce a given time rate of accidents with some average consequence (such as average cost or the number of fatalities, fatal accidents, or fatal plus injury accidents per total accidents). In short, the definition of accident hazard depends on the definition of accident severity, which, in turn, depends on the objective of the highway safety improvement program—whether it is intended to maximize the reduction of total accidents, accident costs, fatalities, fatal accidents, or fatal and injury accidents.

Because the process of relating all dimensional values of the design elements and the design constraints to particular values of hazard is a very complex task, it is extremely difficult to visualize the final product of the synthesis. But, for the sake of illustrating the proposed process of sensitivity analysis, let us assume that the product of the synthesis will take the form of a mathematical model that relates the independent variables that dimension the design elements, the design constraints, and the many interactions thereof. Because of this complexity, the practical application of the synthesis of interactive relations may be highly questionable. Using this kind of formulation for a practical cost-effectiveness decision-making framework may be so cumbersome as to render it useless.

The discussion above suggests that the model be tested for sensitivity to various levels of the independent variables. As the variables that contribute lesser sensitivity are discovered, they are dropped from the model, and the newer, simplified model is tested for predictive precision. This process is repeated, and the least significant variables are successively dropped or combined until the trade-off between predictive precision and simplification for practical application is optimized. The final form of the model will predict a large portion of the variation in the hazard measure by means of the simplest possible model of independent variables.

Successful development of a comprehensive set of design criteria forms the basis for an objective design policy that will enable the highway engineer to design each highway close to optimum. These designs can be accomplished if the art and the science of decision making are placed in the proper perspective, the tools of scientific decision making are brought advantageously to bear at the appropriate points in the design process, and engineering judgment is focused at the appropriate levels. In addition, the comprehensive and objective design policy will provide a framework for assimilating future improvements of design data and technology into the design process.

CONSISTENCY OF GEOMETRIC DESIGN IN RELATION TO DRIVER EXPECTANCY

Consistency has always been recognized as an underlying principle in highway design as exemplified by the following rules of thumb contained in AASHTO design policies. From A Policy on Design Standards (1945):

> Sudden changes between curves of widely different radii or between long tangents and sharp curves should be avoided.

From A Policy on Geometric Design of Rural Highways (1954):

> Horizontal and vertical alignment should not be designed independently. They complement each other and poorly designed combinations can spoil the good points and aggravate the deficiencies of each.

From A Policy on Geometric Design of Rural Highways (1965) (2):

> The 'roller-coaster' or 'hidden-dip' type of profile should be avoided. Such profiles generally occur on relatively straight horizontal alignment where the roadway profile closely follows a rolling natural ground line. Examples of these undesirable profiles are still evident on many highways.

From A Policy on Design of Urban Highways and Arterial Streets (1973):

> Curvature and grade should be in proper balance. Tangent alignment or flat curvature with steep or long grades, and excessive curvature with flat grades, are both poor design. A logical design is a compromise between the two, which offers the most in safety, capacity, ease and uniformity of operation, and pleasing appearance within the practical limits of terrain and area traversed. Wherever feasible the roadway should 'roll with' rather than 'buck' the terrain.

Although the concept of design consistency has been given substantial attention in the design policies, there is a general lack of explicit criteria for the contiguous combination of basic design elements or for the longitudinal variations of such features as horizontal alignment, vertical alignment, and cross section. Without these explicit criteria, highway designers will continue to build inconsistent geometric details into highways.

Recent attention has been focused on design consistency through the development and widespread recognition of the concept of driver expectancy. The general term expectancy relates to a stimulus-response process in which a person with an established set of ideas and concepts is presented a stimulus (visual, auditory, tactile, or other) and responds in some way to this stimulus. Although the stimulus triggers the response, the response may be either directly related or totally unrelated to the stimulus. The person's set of ideas and concepts (predisposition), which greatly influences his or her response to the stimulus, is called expectancy.

Driver expectancy relates to the readiness of the driver to respond to events, situations, or the presentation of information. If an expectancy is met, driver performance tends to be error free. When an expectancy is violated, longer response time and incorrect behavior usually result. Although driver expectancy is similar to the basic expectancy model given above, the expected situation is always changing and environmental factors are more evident, and thus the predictability of the response is reduced. That the response is to an expected situation rather than the actual situation is the vital distinction in understanding the use of driver expectancy in the design process.

APPLICATION OF DESIGN CONSISTENCY

In the most general sense, design consistency means that combination of design elements (and their dimensional specification) that does not violate the abilities of the driver to guide and control the vehicle. Therefore, the concept of driver expectancy is wholly embodied in the general definition of design consistency. In a certain sense, then, the term design consistency can almost be used interchangeably for driver expectancy.

The term driver expectancy relates a subjective appraisal of the adequacy of driver behavioral responses to particular highway situations or conditions. From this general concept is derived the idea of design consistency, which describes those combinations of geometric design elements that do not violate driver expectancies. Thus, human factors engineers, psychologists, highway engineers, and the public for that matter can generally agree that certain extreme combinations of geometric design elements constitute inconsistent design. These are the design features that usually tend to induce noticeable discomfort in the driver.

Using the concept of driver expectancy directly, however, to determine what is or is not consistent design (particularly for those design features that are close to a threshold value) presupposes that driver expectancy can be discretely quantified for a multitude of geometric design configurations. But the feasibility of this kind of quantification is questionable. There do not appear to be any studies that lend quantification (or for that matter even dimension) to the human aspect of driver expectancy. When one looks at driver expectancy as a statistical description of the driving population, a possible basis for quantification might involve observations of overt behavior such as erratic maneuvers. But there are problems in the precision and statistical description of data not to mention the complexity of an experimental design to isolate the effects of design features from the confounding effects of diverse driver, vehicle, and environmental factors. In other words, it is unclear whether it is feasible to isolate the incremental effects of design elements and features on some measure of driver expectancy in empirical studies.

An alternative is to develop criteria (from state-of-the-art syntheses) for performance elements of the driving task based on how critical they are to the safe and efficient operation of individual driver-vehicle components subject to the constraints imposed by geometric design features. In other words, establish time-distance-speed relations appropriate to maintain threshold vehicular stability (both dynamically and in an object-avoidance mode) dependent on the following limitations: driver perception, vehicle performance, driver-vehicle and vehicle-roadway interaction, and combinations of these.

In further describing this approach to evaluating design features for design consistency, it is easiest to talk about the countermeasures to driver guidance and control problems. These may be grouped into at least six general countermeasure approaches:

1. Improve driver detection—These kinds of countermeasures apply mainly to design features that do not fit patterns of driver expectancy and are also difficult to detect but cannot be improved by direct alteration. An example is to change the position of a lane drop that is just over a crest so that it is on an upgrade just downstream from a sag vertical curve.
2. Increase driver perception and response time—These kinds of countermeasures apply to design features that do not fit patterns of driver expectancy and where perception time is limited by sight obstructions. An example is to increase the distance between a crest vertical curve and a close downstream intersection.
3. Eliminate "false cue" designs—These kinds of countermeasures apply to design features that violate driver expectancy and also misguide driver control actions. A prime example is complete redesign to eliminate a side-road intersection that is tangent to a mainline curve.
4. Decrease driver guidance and control demands—These kinds of countermeasures apply to design features that violate driver expectancy in terms of perceiving the critical nature of required speed and path corrections. A typical example is providing a spiral transition to a sharp horizontal curve. Another example is increasing short taper lengths at lane drops or at lane- and shoulder-width transitions.
5. Increase driver expectancy—These kinds of countermeasures apply to design features that violate driver expectancies that are determined by immediately preceding trip experiences. An example here is building in horizontal curvature to "break up" an 8-km (5-mile) tangent section.
6. Build "relief valve" designs—These kinds of countermeasures apply when all other countermeasures are unfeasible. For example, a lane drop can be accomplished by using a painted taper and carrying the full lane width an additional 61 to 121 m (200 to 400 ft) downstream.

Developing this kind of basis requires that performance criteria answer the following kinds of questions:

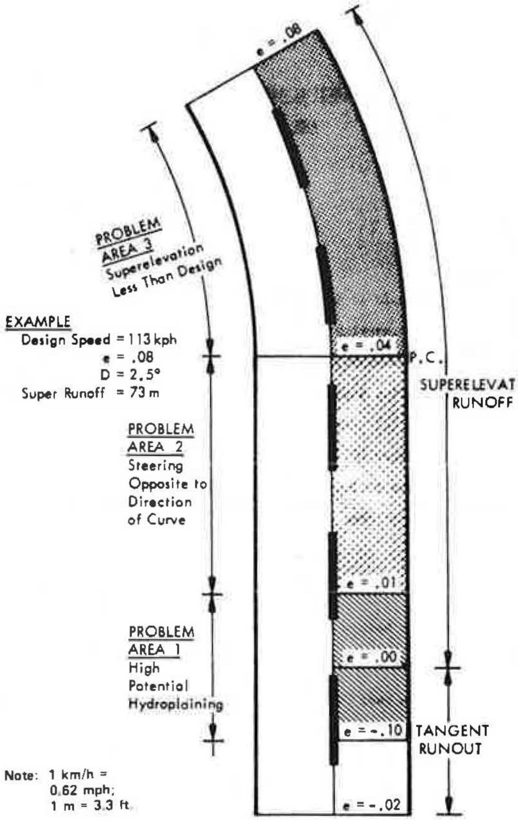
1. What are the threshold values of factors that limit perception of a geometry—e.g., lateral rate of convergence and flat line of vertical sight (parallax)? What is their relation to speed?
2. What are realistic perception times for various design features? Is perception time related to speed?
3. What maximum dynamic response (onset rate of lateral acceleration, braking deceleration, and so on) should be designed for? How does the critical nature of the responses of the driving population relate to various design features?
4. What is the time-distance degradational effect of consistent design on driver expectancy when an inconsistent design feature is introduced?

As a simple and straightforward illustration of this approach, consider the driver traversing from a tangent section to a circular horizontal curve. For horizontal curves, it is standard practice to provide a cross-slope transition from the normal crown on the tangent to full superelevation on the curve. Without a spiral transition, however, this cross-section transition appears to create a compound dilemma. This is most easily illustrated by the example shown in Figure 4. A driver approaching an unspiraled curve is presented first with problem area 1 in which the cross slope is less than 0.01 m/m (ft/ft). Because of this slight cross slope, the pavement does not drain well, and thus a section is created that has a high potential for hydroplaning. The driver no sooner gets through problem area 1 (where he or she may have experienced partial loss of control) than he or she is presented with problem area 2. In problem area 2, the driver may experience some steering difficulty because the cross slope requires steering opposite to the direction of the upcoming curve. When the vehicle passes from problem area 2 to problem area 3, the driver must reverse steering to follow the curve. At this point, if the driver steers the degree of highway curve, the lateral acceleration will be greater than that designed for since problem area 3 does not have full superelevation.

At design speed, for this example, the driver proceeds through the "compound dilemma area" in 2.6 s. Whether a driver can react adequately to these demands on his or her abilities of perception, guidance, and control in the time required is questionable.

If the example is carried one step further, past research (6) shows that drivers do not always expect vehicle stability requirements to be critical on sharper

Figure 4. Cross-slope transition area and related maneuvering problems.



horizontal curves that do not have spiral transitions. As a result, lateral accelerations on sharper curves can exceed assumed design values by as much as 2.13 m/s^2 (7.0 ft/s^2). In addition, because of insufficient space to perform an adequate spiral maneuver (the natural path of the vehicle), the rate of change of lateral acceleration can easily exceed 4.57 m/s^2 (15.0 ft/s^2).

This onset rate is clearly in the range of dynamic instability when one considers the extreme control requirements placed on the driver and the marginal ability of the tire-pavement interface to counteract such extremes.

Why, then, not add spiral transition curves that duplicate the natural path of the vehicle in a noncritical maneuver mode? Even the spiral designs that provide a continuously decreasing radius over the nominal lengths suggested by AASHO (2) will hold the onset rate of lateral acceleration under 0.91 m/s^2 (3.0 ft/s^2)—a rate that is entirely within the stable control range.

The basis of design that is apparent here is that those curves that generate more than some minimum onset rates should have spiral transitions. Although the AASHO "blue book" (2) and most state highway design manuals suggest using spiral transitions, the astute observer is hard-pressed to find a spiral curve in most states. What is apparently needed is a clearer description (sales job) for designers of the critical nature of driver control needs in the absence of the spiral.

DEVELOPING A COST-EFFECTIVENESS METHODOLOGY

Although the development of valid cost-effectiveness evaluation techniques is difficult without the kinds of inputs discussed earlier, these inputs are of little use without the development of an objective cost-effectiveness methodology for the implementation of appropriate design alternatives.

The administrator of a highway department, faced with the task of reducing accident hazard on a jurisdictional basis, must make decisions on the nature of the roadway and desired roadside designs while subject to constraints that affect those decisions. Normally, the principal constraint is limited funds. If there were no funding limitations, certainly the administrator would provide adequate lane widths, large-radius curves, long vertical curves, flat grades, and flat roadsides free of fixed objects close to the roadway. In this situation, the administrator has few decision-making problems. But, in reality, the administrator rarely works with unlimited funds and therefore strives for a strategy that allows the greatest benefits for available funds.

The basis of a cost-effectiveness analysis is that alternative methods are available for reaching an objective and each alternative requires resources and produces benefits. A cost-effectiveness analysis systematically examines the cost and effectiveness (by using some dimensional measures) of alternative methods for accomplishing an objective.

The desired cost-effectiveness methodology requires a complete decision framework for (a) computing the accident hazard associated with any highway location, dependent on the dimensions of its design elements and operational parameters; (b) computing the relative hazard reduction of alternative designs; (c) computing the total cost of a design improvement, including initial costs and differential operational and maintenance costs; (d) computing the relative cost-effectiveness of alternative designs; and (e) choosing the appropriate alternative.

The cost-effectiveness formulation has two basic components, the hazard evaluation and the cost evaluation, as shown below:

$$C/E = C_I/(H_B - H_A) \qquad (1)$$

where

C/E = cost-effectiveness,
C_I = cost of improvement,
H_B = hazard before improvement, and
H_A = hazard after improvement.

As seen in this basic formulation, the hazard evaluation is used twice to compute the hazard reduction. The basic form of the hazard evaluation was discussed earlier. A brief description of the cost evaluation is given below.

In many highway situations, the difference in hazard between design alternatives may be marginal. If this is true, then at least in some cases the cost-effectiveness comparison will be most sensitive to the cost differences between design alternatives. The most important aspect of this sensitivity is the trade-off between the differentials of initial installation costs and maintenance costs. A generalized form of the cost-evaluation model is given below. Although a much more comprehensive form of the model can be anticipated, this example shows the overall concept:

$$C_A = C_I(CRF)_i^t + (C_{MA} - C_{MB}) + [C_{RA}(N_A) - C_{RB}(N_B)] \qquad (2)$$

where

C_A = total net annual cost of design improvement;
C_I = total initial cost of design improvement including costs of design, right of way, removal, grading, paving, and structure and highway-user cost differentials during construction (dollars);

CRF = capital recovery factor;
t = life of design improvement (years);
i = investment return rate (percent/year);
C_{MA} = annual normal maintenance cost after improvement including surface repair and resurfacing, repainting traffic markings, mowing, snow and ice removal, and so on (dollars/year);
C_{MB} = annual normal maintenance cost before improvement (dollars/year);
C_{RA} = annual accident repair costs (to guardrails, bridges, signs, light poles) after improvement (dollars/year);
C_{RB} = annual accident repair costs before improvement (dollars/year);
N_A = annual number of collisions with highway structures after improvement; and
N_B = annual number of collisions with highway structures before improvement.

The cost-effectiveness methodology developed should, by design, lead directly to implementation. The methodology, of course, must be applied within the technical-economic-political decision-making framework of each highway agency, which ranges from the rural township highway department to the state highway department in the most urbanized state. Therefore, the methodology requires a flexible optimization strategy that is responsive to program inputs that vary according to the highway design goals of individual agencies.

The complete cost-effectiveness methodology should have several built-in decision processes other than the basic cost-effectiveness computation. The best use should be made of decision tools such as game theory, linear programming, dynamic programming, control theory, network theory, and various optimization techniques. Furthermore, the methodology should be "compartmentalized" so that subelements can be appropriately altered according to user needs without having to alter the entire methodology.

The development of the description of the design process, which was discussed earlier, will be valuable in identifying many aspects of the complete decision process. These include

1. The integration of the design goals and the highway functional classification into the decision process;
2. The ability to compare design improvements with alternative traffic operational improvements [for example, in many cases the application of "positive guidance" devices (7) may be much more cost-effective than design alternatives such as widening bridges];
3. Incorporation of decision-theory techniques to handle factors of uncertainty;
4. Methods for developing a simpler decision process by using a particular set of solutions of the cost-effectiveness methodology for a given set of input parameters;
5. The ability to optimize "earmarked" improvement programs that are based on subjective decisions or objectives other than safety;
6. The flexibility to accept future refinements in the precision of the interactive relation; and
7. The ability to balance the trade-off between precision and generalization for any particular user.

In relation to item 3—incorporating decision-theory techniques to handle factors of uncertainty—some degree of uncertainty usually affects most of the variables combined in evaluating alternative designs. Sometimes this uncertainty is dealt with by combining "conservative" values. In other cases, the best estimate value is selected for each variable. The decision-theory approach recognizes that the choice has to be made and seeks to structure the problem to incorporate estimates of uncertain factors rather than ignoring them.

In relation to item 7—balancing the trade-off between precision and generalization for any particular user—there is always a trade-off between the degree of precision and the degree of generalization in programming highway safety improvements. Maximum precision requires identifying exact values of all parameters that influence accident hazard. Implementation requires that insignificant parameters be ignored and significant parameters be categorized to minimize the collection of input data. The methodology, therefore, necessarily includes sensitivity analysis at several points in the evaluation. This analysis tests the sensitivity of the decision variable to proposed omissions and generalizations of the input parameters and provides a framework for balancing precision and generalization.

Another integral part of the proposed methodology is outlining the ways and means of implementing the total highway improvement program, as described below.

Implementing the predictive cost-effectiveness program does not mean that a spot-improvement program that uses high-accident-frequency identification procedures should be discarded. Both programs are desirable. The cost-effectiveness program identifies potentially hazardous locations; the high-accident-frequency identification program identifies locations that have demonstrated a high degree of hazard that may or may not be identified in the cost-effectiveness program. Because the cost-effectiveness program cannot precisely account for every single variable that contributes to accident hazard at every particular highway site, certain locations may actually have a higher degree of hazard than that assigned by the cost-effectiveness program. To identify these specific locations, the spot-improvement program may be more appropriate. Then, too, the cost-effectiveness methodology should be helpful in determining the best alternative improvement for sites identified in the spot-improvement program.

Unlike the spot-improvement program, which requires a comprehensive inventory of accident records, the predictive hazard approach requires a comprehensive inventory of site parameters to identify and rank potential improvement sites. Although this inventory could be the most difficult aspect of the implementation program, it may not be as difficult as it first appears. This is where the trade-off between precision and generalization comes into play. The kinds and precision of inventory items should be generalized (simplified) to a degree consistent with the desired level of program precision. It is not necessary to inventory all highways in a jurisdiction before implementing the program. A priority inventory plan can be adopted that accounts for the most sensitive variables in the hazard evaluation. In other words, the inventory plan would assign higher priorities (and hence earlier scheduling) for inventorying high-volume highways, high-speed highways, and high-hazard locations such as intersections.

To determine the general requirements for program funding, statistical procedures can be applied to obtain a representative sample inventory of hazardous locations. By using this sample to generate an estimate for the total population of cost-effective site improvements, the total program funding requirement can be estimated. This indicates to the administrator the general levels of funding that will be needed to meet various program objectives (e.g., the degree of safety payoff over specific periods of time).

SUMMARY

This paper suggests the development of a very comprehensive systems analysis for quantifying the relations between highway design elements (and their combinations) and highway safety. It also suggests the need for developing a rational cost-effectiveness methodology for optimizing the safety payoff of geometric design improvements.

The paper is critical of current AASHTO design policies and, at the same time, is "idealistic" about the potential improvement of these policies. This stance is not intended to sound pretentious but to encourage optimism toward future improvements in the design process. Only by a critical review of current practices can we ever hope to identify the missing links in achieving design consistency. On the other hand, with an idealistic attitude, we can set the highest possible goals for the future, goals that will only be modified by real (and not imaginary) constraints. The antithesis—setting shortsighted goals—would prevent the achievement of solutions that are even close to optimal.

This paper has stressed the need for more sophisticated analysis and decision-making procedures. There is little question of this need for, as the highway community strives more and more for optimality, the tools must necessarily become more objective and complex. This paper, however, does not subscribe to the "black-box" philosophy. The methodology proposed is only a tool and as such must be comprehensible and responsive to the needs of a wide variety of users.

Future design guides must "sell" themselves to the design engineer. Traditionally, the highway design engineer has not directly accounted for the critical nature of the driver's guidance and control needs. The engineer needs to be convinced that this approach to design is not only rational but highly justified. This requires a clear and concise justification of the human-factors criteria that are used in design procedures.

The proposed methodology should be of great value in the design of new facilities as well as in the upgrading or redesign of existing highways and streets. When funds are not available for extensive upgrading of an existing facility, the methodology should aid in demonstrating the cost-effectiveness of upgrading by replacement during normal maintenance procedures. For example, for roadside hazards, as these elements wear out or are damaged or destroyed, they can be replaced by their more cost-effective counterparts.

In these times of increasing litigation against highway departments on the basis of their safety responsibility in highway accidents, a comprehensive and active program of implementing the most cost-effective improvements in order of priority should be a very convincing argument against liability for the government unit. Furthermore, this comprehensive and active implementation program should demonstrate to legislatures and the public the wise use of public funds and thus help avert the predicted deterioration of the existing system because of increasingly inadequate maintenance funds.

REFERENCES

1. Geometric Design Guide for Resurfacing, Restoration and Rehabilitation (R-R-R) of Highways and Streets. AASHTO, 1977.
2. A Policy on Geometric Design of Rural Highways. AASHO, 1965.
3. Manual on Uniform Traffic Control Devices for Streets and Highways. Federal Highway Administration, U.S. Department of Transportation, 1971.
4. A Policy on Criteria for Marking and Signing No-Passing Zones on Two and Three Lane Roads. AASHO, 1940.
5. D. W. Harwood and J. C. Glennon. Framework for Design and Operation of Passing Zones on Two-Lane Highways. TRB, Transportation Research Record 601, 1977, pp. 45-50.
6. J. C. Glennon and G. D. Weaver. Highway Curve Design for Safe Vehicle Operations. HRB, Highway Research Record 390, 1972.
7. G. J. Alexander and H. Lunenfeld. Positive Guidance in Traffic Control. Federal Highway Administration, U.S. Department of Transportation, April 1975.
8. J. C. Glennon. The Desirability of Highway Designs for Operating Speeds Above 70 mph. 40th Annual Meeting, ITE, Houston, TX, Sept. 1970.
9. D. R. Drew. Traffic Flow Theory and Control. McGraw-Hill, New York, 1968.
10. D. L. Woods. Application of Driver Expectancy in Design. Texas Transportation Institute, 1972.

Discussion

Sally Free, Center for Auto Safety, Washington, D.C.

The Center for Auto Safety agrees that the procedures now used to formulate highway design standards are not only inadequate and obsolete but also detrimental to the safety of the traveling public. Present standard-setting methodologies have failed to reduce the annual highway death toll below the staggering 40 000 mark. It is apparent that radical decreases in accident and fatality rates can no longer be realized by simply applying "common sense" solutions to old problems.

The system now in effect is one of flexible standards—that is, standards that are subject to negotiation between the Federal Highway Administration (FHWA) and the states. Vague terminology, so-called engineering judgment, and exhortatory language are poor substitutes for clear, mandatory performance criteria and objectives.

Instead of independently establishing objective performance criteria to ensure the safe design and construction of the nation's roads, FHWA has simply incorporated by reference many AASHTO design policies. One need only look at the Code of Federal Regulations, Volume 23, Part 625, to see the influence of AASHTO. The close involvement between FHWA and AASHTO during all stages of the decision-making process has had a tremendous effect on how agency decisions are made—important decisions that affect the public's safety, pocketbook, health, and environment. Indeed, this unique partnership has allowed AASHTO to shape the direction of federal standards and policies, the content of rule-making, and even the enforcement capabilities of the regulating agency—FHWA.

This represents a rather disturbing situation since the state highway departments as a result have remained largely self-regulated, writing their own standards through AASHTO. These standards are far from optimal; much of the time the needs and safety of the driver are neglected in favor of wording that is excessively advantageous to and protective of highway officials and departments. The principal motives and objectives behind many of these design standards are the hopes of highway officials that the policies will lessen liability, decrease costs, and increase state discretionary use of federal money. These standards and policies are often simply a collection of suggestions and recommendations that are lacking in detail, are highly qualified and ambiguous, and provide so-called technical guidelines that have not always been substantiated by research or field experience. In reference to the imprecise use of terminology, one FHWA attorney has noted that "the 'standards' are so easily circumvented that they have become essentially meaningless."

The National Transportation Safety Board (NTSB) has characterized highway design standards as having originated from a "fragmented remedial approach" to safety. That is, isolated safety improvements are made based on "post mortem investigations" rather than initiating a systems approach to accident prevention at the design stage. The 1969 NTSB study, Compatibility of Standards for Drivers, Vehicles, and Highways, points to everyday traffic situations that illustrate the interrelationships of all elements in the highway system. The study maintains that the highway community has not adequately considered these interrelationships in the development and issuance of highway design standards. As a consequence, many design standards for different surface transportation subsystems are incompatible, and highway operating and design problems result. The development of performance-based design standards accompanied by a rational explanation of the function of these standards is recommended.

Although the concept of systems compatibility or systems engineering has been advanced since the 1960s, there has been little acceptance of the idea by highway departments. This lack of success stems not from a faulty methodology but rather from the need to change old policies and attitudes. First and foremost among the policy changes needed as a condition for the implementation of an effective systems engineering approach is for FHWA to establish itself firmly as a regulator and promulgate its own standards. FHWA can no longer be simply a mechanism for resource transfers. In addition, FHWA and other standard-setting agencies such as the National Highway Traffic Safety Administration (NHTSA) must better coordinate and communicate their policies, standards, and rulemaking procedures. A responsible approach to improving the process must include an extensive review and evaluation of all standards to determine their relevance and compatibility. For systems engineering to succeed, design engineers must begin thinking in holistic terms about the highway environment rather than just reacting to isolated problems that surface as a result of the present piecemeal design approach.

The systems engineering approach is advantageous in that it forces highway engineers to think in new and different terms. Isolated improvements and short-term solutions as a basic approach may be shown to be the least cost-effective alternatives over the long term. Installation expenses, maintenance costs, safety benefits, and operational efficiency can be more meaningfully evaluated when the transportation system is viewed and designed as a functional whole. An understanding of the total system operation will enable engineers to better identify and predict problems, evaluate alternatives, and implement solutions.

Systems engineering should help reduce tort liability. In recent years, the willingness of the courts to hold highway agencies and officials accountable for faulty highway design has caused tort liability to be of major concern. In an attempt to justify highway deficiencies and faulty design as accepted practice, many highway agencies are urging the adoption of lower standards. It is their hope that courts will no longer hold the highway agencies accountable to the higher standards and that this will substantially reduce their exposure to tort liability.

This attempt is ill-advised and legally misguided. The cause of liability suits is not the standards but hazardous roadway conditions. A lowering of standards can only serve to increase fatalities and injuries and thereby correspondingly increase the number of claims made against highway departments. Incompatible and inadequate standards will give the lawyer the opportunity to "pick and choose" the standard that best suits the needs of a client. At a conference session on the compatibility of standards at the Fifty-sixth Annual Meeting of the Transportation Research Board, FHWA trial lawyer David Oliver reached the following conclusion:

> Without standards, accidents will occur and legal judgments will ensue. Without compatibility, standards will be not only unenforceable but also indefensible. Without cooperation there will be no 'standards.' The driver, vehicle, highway design functions must be integrated or the legal function will bare its teeth.

C. William Gray, Ohio Department of Transportation

The subject of this paper is timely, and its purpose—to promote a more systematic approach to highway design—will be enthusiastically supported by highway designers when the concept has been developed to a usable level. I have read the paper from the outlook of a designer instead of that of a researcher, since design is my background, and I believe this paper will have little impact on design until more research is performed and the system is much more thoroughly developed. The urgent need and motivation for design policy changes are clearly and accurately stated in the introduction in the statement that we are rapidly changing from a massive road-building campaign to one of improving the traffic safety and service of existing highways.

Those of us who have studied AASHTO design policies and applied them to highway design consider them to be excellent publications. If the use of AASHTO design policies has failed to meet the needs of today's drivers, perhaps the blame rests with the people who have not used these policies as a basic foundation for design and then added to that foundation from the vast store of information available from operational data and experience and current research findings. That is a very difficult task in today's rapidly changing world, and I think that is really what this paper is trying to do.

I might observe here that even our language is rapidly changing and that perhaps it does not need to change so much. Practicing highway design engineers would more readily understand and adopt new concepts and design policies if they were expressed in more commonly understood words. The last statement of the introduction is an example of how words can be hard to understand. In discussing design for the needs of the driver, it says, "But this goal can only be accomplished if the design process is objectified to the extent that it maximizes the effectiveness of design improvements subject to funding constraints." I think that says spend your money where it will do the most good. Having said that either way, we now need to go much beyond what this paper does to explain to the highway designer how to maximize design effectiveness or how to do the most good with our money.

The paper does recognize communication problems by stating the following in the summary: "The methodology . . . must be comprehensible and responsive to the needs of a wide variety of users." We could also say that the policies must be understood by highway designers so that they can apply them to all types of highway projects.

The paper recognizes, in its discussion of the evolution of highway design, the recent application to design of the systems engineering approach, which uses game theory, queuing theory, linear programming, dynamic programming, control theory, critical path methods, network theory, and various optimization techniques. Just the statement that all those things have recently evolved pinpoints the difficulty designers have in keeping current. I do not really understand some of these techniques, and I believe that many designers share my view. (Note that none of the theories named are included in the list of references at the end of the paper.)

Harwood and Glennon's discussion of objectifying the design process is based on Figure 1 but does not convey a clear understanding of the figure. At that point, it is clear that the paper will not provide a designer with an objective design process to use today, but, as the paper states, a major research effort is needed to produce a design process that will achieve safe, consistent highway designs.

The statement that "the only design feature that physically governs maximum safe speed (for modern highway designs) is horizontal curvature" should be modified to add sight distance. Maybe only horizontal curvature governs maximum speed but, in considering maximum safe speed, sight distance is a very important design feature and must be included with horizontal curvature as a governing feature.

Figure 3 shows an involved concept—a matrix for the synthesis of interactive relations—without a clear explanation.

The discussions of the consistency of geometric design in relation to driver expectancy and application of design consistency are appropriate. These subjects are of much greater concern to designers today than they were a decade ago, and they should be a major influence in future design policies.

In discussing the cost-effectiveness methodology, the authors have recognized the realistic nature of highway improvements by stating the following: "The methodology, of course, must be applied within the technical-economic-political decision-making framework of each highway agency. . . ." That has been true in the past and I am sure it will continue to be true in the future.

In view of the ever-increasing demand to improve our highways for greater traffic service and safety, it is a necessity that the highway designer have cost-effective design decision tools as proposed in this paper. I hope the paper will result in subsequent research and progress toward an early achievement of usable modern design policies, and I would encourage the authors and others to continue to work toward that objective.

Authors' Closure

We want to thank both Free and Gray for their discussions. Both of their viewpoints—Free's as a highway safety advocate and Gray's as a state highway designer—are different from our own, but their discussions help to both clarify and add depth to the intent of our paper.

Much of Free's discussion highlights the points made in our paper. She says, "The development of performance-based design standards accompanied by a rational explanation of the function of these standards is recommended," and we agree. She says, "Isolated safety improvements are made based on 'post mortem investigations' rather than initiating a systems approach to accident prevention at the design stage," and we agree. She says, "Systems engineering should help reduce tort liability," and we agree. She also says, "Installation expenses, maintenance costs, safety benefits, and operational efficiency can be more meaningfully evaluated when the transportation system is viewed and designed as a functional whole," and again we agree.

Although almost half of Free's discussion is in tune with our technical thesis, the other half gets into a far-reaching indictment of the process of setting national design standards. We definitely disagree with Free's opinion that the FHWA-AASHTO partnership has been some sort of back-room conspiracy aimed at protecting some vested interests of the state highway agencies at the expense of the motoring public. Both FHWA and AASHTO obviously share Free's deep concern for highway safety because they have (independently and jointly) sponsored many of the technological developments that have led to the improved safety performance of our highways. The roles of FHWA and the states as the major supporters of the Transportation Research Board belie Free's argument.

Free states that present standard-setting methodologies have failed to reduce the annual highway death toll below the staggering 40 000 figure. Our question is, Who ever deduced that highway design practices are the major contributor to highway accidents? Then, too, how can "present standard-setting practices" themselves ever make a measurable impact without the political recognition of the very large funding allocations needed? We must keep in perspective that it is difficult to change quickly the momentum created by hundreds of thousands of kilometers of highway that were designed and built before the advent of modern highway design technology.

What we have attempted in our paper is not to suggest discarding the present methodologies that Free claims are obsolete and detrimental but rather to recognize that the scenario of highway development has changed dramatically and that it is time to fine-tune our design methodologies so that highway agencies can reach a better balance between their safety responsibility and their fiscal responsibility. This balance cannot be achieved by using Free's "more is better" philosophy. This is why cost-effectiveness analysis is important to the design process. Although many highways could justify even high-cost safety improvements, there are other highways—particularly in the category of low-volume local roads—that cannot justify any safety improvements at all.

Gray has also highlighted many of the points in our paper but from a different perspective than Free's or ours. In addition, Gray's discussion is more a direct critique of the paper. We welcome his homespun language. He has rightly pointed to our flaws in clearly communicating our thesis. Communication is a constant problem in any profession. Idea papers often go for years without being understood by practitioners. This is mostly the fault of the authors, but then, too, the process of adapting abstract concepts so that they fit concrete and practical applications is naturally difficult and always requires considerable input by the practitioner, who usually is not paid to deal with concepts.

We also appreciate Gray's confession that he knows very little about most of the established systems engineering tools of optimization. This, of course, does not reflect on his stature as a highly respected member of the highway engineering community. Gray's statement, however, does raise a question: Why is highway design one of the few engineering professions that does not know of and regularly use these systems techniques?

In closing, we again thank the discussants for their responsive inputs. It was exactly this kind of open dialogue that we were trying to generate. Our only hope is that the dialogue will continue.

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