Intergreen Interval Controversy: Toward a Common Framework

C. S. PAPACOSTAS AND NEAL H. KASAMOTO

The chronological development of the most commonly used intergreen interval formulas is traced and a disparity is disclosed between the interpretation presumed by ITE and that originally proposed by Gazis et al. A realistic example clearly shows that proper application of the speed-location diagram introduced by Gazis et al. can enhance the traffic engineer's judgment and can provide a consistent means of reporting research-related observations. The speed-location diagram of the intergreen interval problem must be adopted as a standard tool by traffic engineering practitioners and researchers.

The intergreen interval, that is, the total time period between conflicting green displays at signalized intersections, is the subject of intense debate among traffic engineers. The intergreen interval is commonly displayed either as a steady yellow interval or a combination of yellow followed by an all-red period. An atypical method was used in Pittsburgh, Pennsylvania, and in Ketchikan, Alaska, at least into the late 1960s (*1*,*2*). In that case, the intergreen interval was displayed as a sequence of a simultaneous green and yellow interval followed by a standard yellow interval.

As far back as 1929, Matson, though viewing the intergreen interval as merely an intersection clearance period, wrote that "there are many ways of indicating this caution or clearance period. . . . An understanding of the effects of the clearance period is essential in determining just what is needed. When a definite statement is made as to what amount of time shall be set aside for clearance periods in each cycle, the choice of how these periods shall be indicated rests with the public and its education" (*3*). Yet, after more than 60 years, no consensus has emerged relative to any of these requirements. In 1989, the ITE Technical Council Committee 4A-16, having conducted a review of the vast literature on the subject, proposed revisions to its recommended practice in which it acknowledged that "[d]ivergent and strongly held positions are common when engineers discuss vehicle change intervals. . . . Even among engineers who agree on the method, there are disagreements relative to application" (*4*).

This paper presents an independent review of the chronological development of intergreen interval design equations, discusses the major differences between them, and shows that there are two disparate interpretations of the common design equation that is based on the equations of motion: the interpretation implicit in the *ITE Handbook* and that proposed by Gazis et al. (*5*). This paper shows that the correct interpretation of the Gazis et al. proposal provides a general framework that can help unify what may first appear to be irreconcilable differences of opinion. When properly used, speed-location diagrams in the form suggested by Gazis et al. can enhance the engineer's judgment of intergreen timing at specific intersections and can also aid researchers in properly reporting and interpreting their empirical data.

Regarding terminology, several terms have been used in the literature to refer to the intergreen interval and its subdivisions. Some of these terms (e.g., "clearance interval") attempt to convey a description of purpose or function, but disagreement about these terms has caused unnecessary communication difficulties. ITE (*4*) currently uses the term "change interval," a sufficiently neutral term but one that has also been used to refer to only the yellow display (*6*). The term "intergreen interval," which refers to the total time between conflicting green phases, was borrowed from Hulscher (*7*). The notation used in this paper is partly the authors'. The use of subscripts to certain variables encountered in intergreen interval formulas is an attempt to emphasize the differences in interpretation given to these variables by different authors. A significant part of the controversy regarding the timing and display of the intergreen interval will be traced to these differences.

EVOLUTION OF DESIGN EQUATION

Matson Model

In 1929, Matson (*3*) proposed a formula for computing the needed "clearance interval" to allow vehicles crossing the stop line at the onset of this interval to clear the width of the intersection (W) before control is transferred to the cross street. He also used the terms "amber period" and "caution period" to describe the subject interval. Matson's primary concern was the proper timing and coordination of fixed-signal systems to accommodate the progression of traffic waves traversing urban streets. The required duration was simply taken to be equal to the intersection width divided by the "speed which is normal to the area traversed," that is,

$$T = W/V_n \tag{1}$$

In his short treatment of the subject, Matson described the elaborate procedures needed to establish the "approximate speeds which will be suitable for a signalized street."

C. S. Papacostas, Department of Civil Engineering, University of Hawaii at Manoa, 2540 Dole St., Honolulu, Hawaii 96822. N. H. Kasamoto, Hawaii Department of Transportation, 79 South Nimitz Highway, Honolulu, Hawaii 96813.

1950 ITE Handbook

The 1950 edition of the *Traffic Engineering Handbook* (*8*) used the terms "clearance interval" and "yellow signal indication" and suggested adding the "minimum driver stopping distance" (S_{min}) to the numerator of Equation 1, yielding

$$T = (W + S_{min})/V_n \quad (2)$$

The rationale for adding the stopping distance was that a vehicle traveling at the "normal intersection approach speed" (V_n) could either stop (if located farther than S_{min} from the stop line at the onset of yellow) or clear the intersection at a constant speed (if located closer than S_{min} from the stop line at the onset of yellow).

The following formula, attributed to Earl Reeder, then director of Traffic and Transportation for the city of Miami, Florida, was also presented:

$$T = 0.8 + 0.04V_n + 0.7W/V_n \quad (3)$$

where V_n is given in mph and T in seconds.

Equation 3 results from substituting in Equation 2 the traditional stopping distance formula based on an equivalent constant deceleration rate, that is,

$$S = tV + V^2/(2d) \quad (4)$$

where t is the perception-reaction time and d is the deceleration rate. Apparently, Reeder used a perception-reaction time of 0.75 sec (rounded up to 0.8) and a deceleration rate of 17 ft/sec^2 along with conversion factors allowing the specification of V_n in mph and W in feet. The basis for these values is found in another section of the handbook, "Stopping Distances Used for Design Purposes." The following statements are also found in the 1950 handbook (*8*, p. 69): "Deceleration considered undesirable but not alarming to passengers is 11 feet per second per second," and "comfortable deceleration is 8.5 to 9 feet per second per second." Thus, the stopping distance implicit in the Reeder formula represents emergency rather than comfortable conditions. Moreover, it is not concerned with the deceleration rate that would be attainable on *wet* roadway surfaces, which is the condition governing design in various aspects of highway and traffic engineering. Parenthetically, the traditional design expression of deceleration in terms of kinetics is given by:

$$d = g(f \pm G) \quad (5)$$

where

g = acceleration due to gravity,
f = equivalent coefficient of friction representative of the overall speed change, and
G = roadway gradient.

A friction coefficient of about 0.3 (with some variation related to initial speed) is generally suggested as an appropriate value in calculating safe stopping distances on wet pavements. For a level or nearly level roadway, this value of f leads to an equivalent constant deceleration rate of about 10 ft/sec^2, which happens to equal the widely reported value of comfortable deceleration on dry pavements. The purely kinetic Equation 5 should be preferable to the mixed kinematic-kinetic formula suggested by Parsonson and Santiago (*9*) and adopted by ITE (*4*) because it makes explicit the effect of friction on safe operation. The choice of a high design value for deceleration by Reeder and others was apparently motivated by a desire to keep the duration of the change interval low in order to satisfy those practitioners who "frown on the idea of using yellow intervals in excess of 3 to 5 seconds" (*8*, p. 226).

Substitution of Equation 4 into Equation 2 yields the following general formula:

$$T = t + V_n/(2d_e) + W/V_n \quad (6)$$

where d_e is emergency deceleration.

The 1950 handbook also raised the possibility of deducting from the calculated change interval duration the time required by the leading stopped cross-street vehicle to accelerate from its stop-line position to the point of conflict with the clearing stream. The rationale for this deduction was also discussed much later (in 1977) by Williams (*10*), who, nevertheless, warned that the "time deduction for cross-flow acceleration needs to be applied with caution, and a value of zero should be used if light jumping is possible." In 1981, Parsonson and Santiago (*9*) also pointed out that "the concept pertains to stopped traffic starting up on the green, and not the vehicles approaching the intersection at speed when their signal goes green."

Matson et al.

A restatement of Equation 6 appeared in the Matson et al. (*11*) discussion of the needed yellow light period, except that they prescribed using the comfortable rather than emergency stopping distance. In an obscure theoretical derivation, they computed and compared the required time to stop (y_1) and the required time to clear (y_2) at constant vehicle speed a distance equal the stopping distance plus the intersection width. Cast in different notation than theirs, these two times are:

$$y_1 = t + V/d^* \quad (7)$$

$$y_2 = t + V/(2d^*) + W/V \quad (8)$$

where d^* is comfortable deceleration rate.

From the general comparison of y_1 and y_2, Matson et al. concluded that "time to stop becomes the critical value in determination of yellow light at higher speeds, though time to clear may be the critical value at lower speeds." However, they gave no explanation as to why they felt that the time to stop, as interpreted above, should be used as a criterion for setting the intergreen interval period. As an ITE committee pointed out much later, "once a driver decides to stop, the displayed signal indication becomes meaningless" (*12*). Nevertheless, the concept entered the consciousness of many traffic engineers and, without a doubt, has slanted their understanding and interpretation of the problem.

Gazis et al.

Gazis et al. (*5*) took a fresh look at what they called "the problem of the amber signal light" and formulated an analytical model to describe the predicament of a driver approaching a signalized intersection at the onset of yellow. They essentially expressed the uniform-acceleration equations describing the two possible maneuvers available to the driver, that is, either decelerating to a stop or attempting to clear the intersection, accelerating if necessary. Equation 9 gives the minimum stopping distance from which a vehicle traveling at an initial speed (V_o) can come to a comfortable stop, that is, at a *comfortable* deceleration rate (d^*):

$$X_s = t_s V_o + V_o^2/(2d^*) \tag{9}$$

where

X_s = minimum stopping distance,
t_s = perception-reaction time associated with the decision to stop,
V_o = vehicle speed at the onset of yellow, and
d^* = comfortable deceleration.

The parabola of Equation 9 is independent of the duration of the intergreen interval.

The maximum distance (X_a) from which a vehicle traveling at an initial speed (V_o) can *just* clear the intersection of width W during the intergreen interval of duration T, accelerating if necessary, is given by

$$X_a = V_o T + (1/2)a(T - t_a)^2 - (W + L) \tag{10}$$

where

X_a = maximum clearing distance,
a = equivalent constant acceleration rate,
T = duration of change interval,
t_a = perception-reaction time associated with the decision to clear the intersection,
W = intersection width, and
L = vehicle length.

Gazis et al. reasoned that if, for whatever reason, a vehicle cannot accelerate beyond its approach speed, Equation 10 becomes

$$X_0 = V_o T - (W + L) \tag{11}$$

This linear equation has an X_0-intercept of minus $(W + L)$ and a slope equal to T.

Reasoning that a vehicle approaching at the speed limit (V_1) should not be expected to accelerate in order to clear the intersection, Gazis et al. proposed that the speed limit be used for design purposes. Under this assumption, they proposed that the minimum duration of the change interval be given by the solution of Equations 9 and 11 after setting $X_S = X_0$, that is,

$$T_{min} = t_S + V_{des}/(2d^*) + (W + L)/V_{des} \tag{12}$$

where $V_{des} = V_1$ = the "design" speed used to calculate T_{min}.

For purposes of discussion, Figure 1 plots X_S and X_0 as functions of individual-vehicle approach speed V_o (i.e., Equations 9 and 11). When the two lines intersect, as in the case shown, the speed-position space is divided into five regions as follows:

A: A vehicle cannot clear at constant speed but can stop comfortably,
B: A vehicle cannot stop comfortably but can clear at constant speed,
C: A vehicle has the option to execute either maneuver,
D: A fast-moving vehicle can execute neither maneuver, and
E: A slow-moving vehicle can execute neither maneuver.

Gazis et al. introduced the term "dilemma zone" to describe the conditions encompassed by Regions D and E. A dilemma zone is said to exist if, for a given approach speed, $X_S > X_0$, as illustrated in connection with V_3 and V_4 in Figure 1. The length of the dilemma zone is given by the corresponding differences $(X_S - X_0)$ as shown. A vehicle traveling at a speed associated with a dilemma zone, however, will experience the problem only if it happens to be located within the dilemma zone at the onset of the intergreen interval. An approaching vehicle at the same speed, V_3 or V_4, would either be able to stop or be able to clear the intersection without accelerating if located in Regions A or B, respectively. Figure 2 shows a situation where the X_S and X_0 lines are tangent to each other, that is, where the dilemma zone region is eliminated for only one value of speed. Figure 3 shows a situation where the X_S

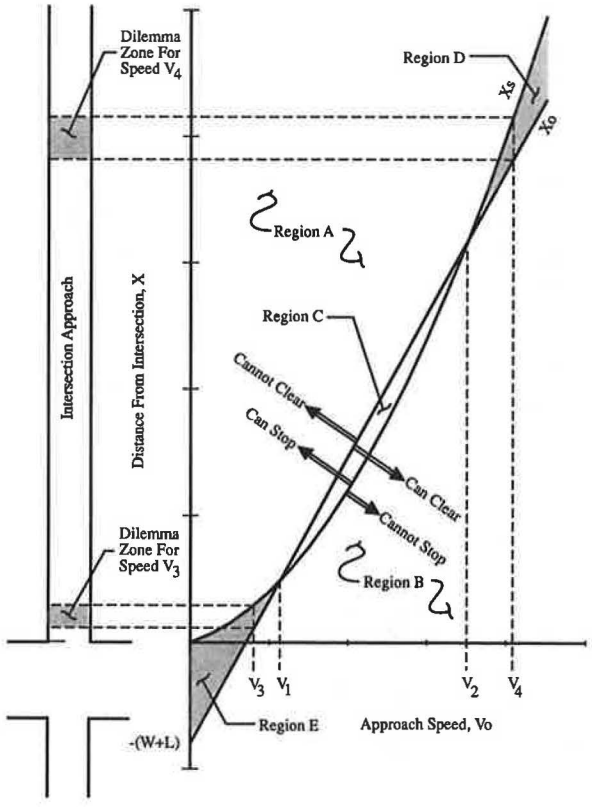

FIGURE 1 Case of intersecting X_0 and X_S plots.

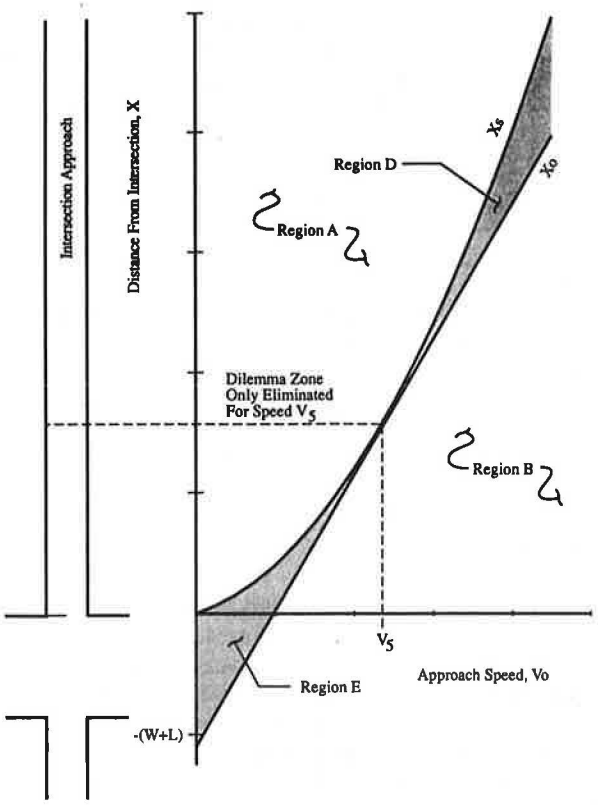

FIGURE 2 Case of tangent X_0 and X_S plots.

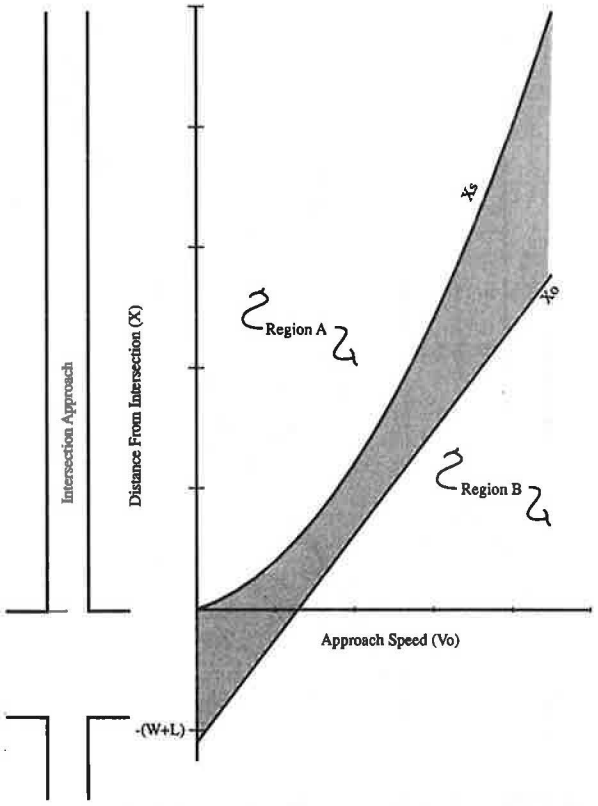

FIGURE 3 Case of nonintersecting X_0 and X_S plots.

and X_0 lines do not intersect and a dilemma region is associated with all speeds. This situation arises when T is set at a value smaller than the T_{min} value obtained by Equation 12. Again, even though dilemma zones may be encountered at all approaching speeds, not all approaching vehicles would in fact be located within the dilemma region at the onset of the intergreen interval.

It was clearly not the intent of Gazis et al. to suggest that the T_{min} value calculated by Equation 12 would eliminate the dilemma zone problem for all approaching speeds. In fact, they discussed the implications that the intergreen interval duration determined by Equation 12 would have on vehicles approaching at speeds other than V_{des}. They also examined the case of accelerating vehicles according to a nonconstant acceleration-speed relationship of the form $a = A - BV$. The presence of dilemma and option regions with and without acceleration are shown in Figure 4 for a case similar to that described by Figure 3. In general, if acceleration is possible, option zones tend to appear over a longer range of individual-vehicle approach speeds.

The presence of option zones (i.e., Region C) did not appear problematic to Gazis et al. However, in 1981, Bissell and Warren (*13*) pointed out that excessively long option zones may contribute to rear-end collisions when the driver of a leading vehicle chooses to stop and the driver of a following vehicle in the option zone decides to go.

Finally, it is important to note on the typical speed-location diagram the presence of a triangular area below the speed axis corresponding to low speeds. This area describes the situation where a vehicle already in the intersection area at the onset of the intergreen interval would not be able to clear the intersection at constant speed. A dilemma zone region also exists to the right of this triangle for slow-moving vehicles

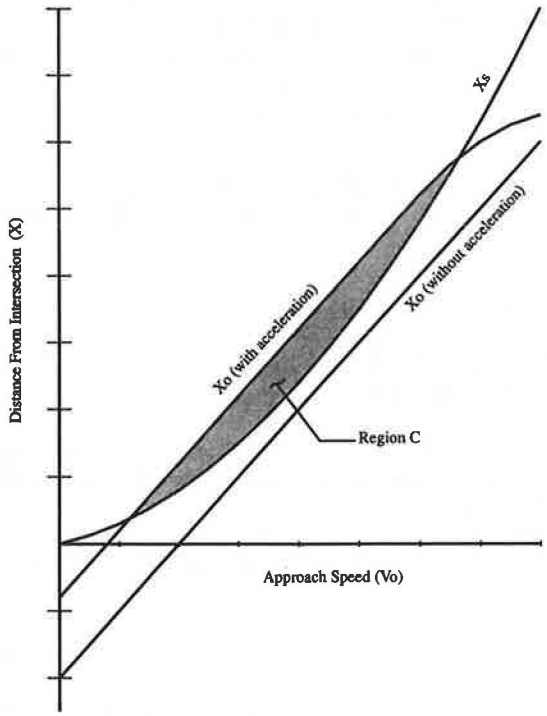

FIGURE 4 Effect of acceleration, when possible.

located very close to the stopping line at the onset of the intergreen interval. This region can be reduced but cannot be completely eliminated, no matter how long the duration of the intergreen interval is. The likelihood that approaching vehicles would actually experience the described conditions would be higher during periods of congested flow when speeds are low and acceleration is restricted by vehicles ahead. Unfortunately Gazis et al. failed to adequately emphasize this situation.

1965 ITE Handbook

The 1965 edition of the Traffic Engineering Handbook (*14*) presented two equations for the determination of the yellow interval as follows:

$$y_1 = t + V_a/(2d) \quad (13)$$

$$y_2 = t + V_a/(2d) + (W + L)/V_a \quad (14)$$

where V_a is specified generically as the "approach" speed. The handbook stated that Equation 13 yields the "minimum time to stop," whereas Equation 14 (attributed to Gazis et al.) yields "the minimum time to stop or clear the intersection." The 1965 handbook suggested that (*14*)

> The yellow clearance interval used should exceed the values of y_1 for the approach speed selected. . . . Where y_2 exceeds the value selected for the yellow interval and where hazardous conflict is likely, an all-red clearance interval is frequently used between the yellow interval and the green interval for opposing traffic.

The use of these two equations was apparently motivated by the work of Matson et al. (*11*), as reflected in the pair of Equations 7 and 8. Inexplicably, however, Equation 13 is misspecified because it does not yield the time to stop. If needed, the proper equation for this time to stop would be Equation 7. Nonetheless, this error has persisted over the years. As late as 1986, Wortman and Fox (*15*) interpreted Equation 13 (and, therefore, the first two terms on Equation 14) as representing "the time required for the driver to come to a safe stop." Lin's work (*16*) was also predicated on the same misconception. His explanation of Equation 14 is as follows:

> The sum of the first two terms in this equation represents the time required for a driver to come to a stop after the yellow interval begins. The last term of the equation is the time required to cross the intersection.

This unfortunate error has undoubtedly caused considerable confusion regarding the interpretation of Equation 14. It appears that some of the critics of the kinematic model have unknowingly leveled their criticism on their misinterpretation of the model rather than on the model itself.

1976 ITE Handbook

The 1976 edition of the *Transportation and Traffic Engineering Handbook* (*17*) retained Equations 13 and 14 and attempted to discuss the dilemma zone concept but attributed it to Olson and Rothery (*18*) rather than Gazis et al. (*5*). In this connection it stated that:

> An incorrect choice for the length of yellow period . . . can lead to the creation of a *dilemma zone*. This is an area close to an intersection in which a vehicle can neither stop safely nor can clear the intersection before the beginning of the red interval without speeding.

Equation 14 is referred to as yielding the "non-dilemma yellow period." The terseness of the dilemma zone description in the handbook and the fact that the citation of Gazis et al. (who had fully presented the concept) was dropped may have caused additional confusion to some users of the handbook. Specifically, some users, unfamiliar with the Gazis formulation, could have been left with the erroneous impression that the choice of a yellow interval of y_2 is meant to eliminate a dilemma zone for all approaching vehicles. As Figures 1 to 3 show, this is not the proper interpretation of Equation 14.

Additional difficulties may have been introduced by the way in which the 1976 handbook describes, as quoted above, the dilemma zone and its relation to the length of the intergreen interval. As Figure 1 shows, dilemma zones can exist close to the intersection for vehicles approaching at slow speeds (for example, V_3, Figure 1), and farther away from the intersection for fast-moving vehicles (for example V_4, Figure 1), even in instances where the length of the intergreen interval is set at y_2 according to Equation 14. Moreover, Figure 2 shows a situation where dilemma zones exist for some vehicles approaching at all speeds other than that used to calculate y_2. As mentioned before and contrary to the above quote, the possibility that slow-moving vehicles could encounter dilemma zones very close to the stop line cannot be eliminated altogether.

Another source of confusion is the 1976 handbook's implication that individual drivers located in a dilemma zone close to the intersection can clear the intersection by *speeding*. It is true, and Gazis et al. addressed this question, that under certain (but not all) conditions, drivers in a dilemma zone at the onset of the intergreen interval can clear the intersection by *accelerating* (see Figure 4). Exceeding the speed limit (i.e., speeding) is not a prerequisite to clearing the intersection under all such circumstances. The likelihood of drivers having to exceed the speed limit in order to clear the intersection before the beginning of the red interval is higher when the traffic is light and vehicles are approaching at relatively high speeds.

1982 ITE Handbook

Under the heading "yellow change and clearance intervals," the 1982 edition of the *Transportation and Traffic Engineering Handbook* (*19*) reintroduced a reference to Gazis et al. and explained their rationale relating to the dilemma zone problem. In an apparent attempt to rectify the misspecification of Equation 13 described above, the handbook begins with the stopping distance equation, that is,

$$S = tV + V^2/(2d) \quad (15)$$

It divides both sides by V so that the left side becomes identical

to that of Equation 13. The handbook calls the resulting term S/V the "minimum clearance time" needed for a "driver to proceed *into* the intersection." Thus, what in earlier editions was considered to be the "time to stop" is now given a different interpretation, that is, the minimum "clearance time" required by a vehicle traveling at a constant speed (V) to cover the stopping distance (S) that would be traversed if the vehicle were to come to a stop from its initial speed (V). In the context of a design formula where S is set to a particular value of S_{min}, this so-called "clearance time" applies only to a nonaccelerating vehicle traveling at the speed used for design purposes (V_{des}). Whether such a vehicle would in fact reach the stop line after S_{min}/V_{des} seconds depends on its initial location at the onset of the intergreen interval.

The value obtained via Equation 14 is described in the handbook as the minimum clearance time that would permit a driver to proceed *through* the intersection. As with the preceding case, the handbook does not make clear the fact that this condition would be satisfied only by vehicles approaching at the speed used in Equation 14 to calculate the intergreen interval requirement that happen to be located at or closer than the comfortable stopping distance implied by that speed.

True Gazis Contribution

Despite references in the ITE handbooks to the work by Gazis et al., the ITE conception of the dilemma zone problem has remained faithful to Matson's original idea of the clearance interval, that is, the time required by vehicles traveling at a single "control" speed to traverse the comfortable stopping distance (S_{min}) corresponding to that speed plus the width of the cross street. According to this restricted conception, vehicles that happen to be located closer than S_{min} at the onset of the intergreen interval would be able to clear the intersection before the cross street receives a green signal, whereas vehicles located farther away would be able to stop comfortably. In other words, ITE's failure to explicitly assess the implications of setting the duration of the intergreen interval, according to Equation 14, on vehicles that approach the intersection at speeds other than that used for design has led ITE to strongly imply that the duration given by Equation 14 can totally eliminate the dilemma zone problem. In terms of their mathematical form, the basic ITE formula (Equation 14) and the Gazis formula (Equation 12) are identical. However, the greatest contribution of Gazis et al. to the understanding of the problem is not that they came up with a formula that had been around for a long time. Rather, their main and, unfortunately, least appreciated contribution lies in the fact that they presented the larger framework, the speed-location diagram shown in Figures 1 through 4, which must *always* be used in judging the appropriateness of the calculated intergreen interval. Thus, given the speed-location diagram of Figure 2, only an imprudent engineer would accept the intergreen duration shown for a design speed that happens to be equal to V_5 merely because it satisfies Equations 12 and 14. Without reference to the speed-location diagram, the presence of a dilemma zone for all speeds other than V_5 would not be readily evident. Similarly, the value of the intergreen interval corresponding to design speeds V_1 and V_2 in Figure 1 also satisfies Equations 12 and 14. However, the dilemma and option implications are distinctly different in the two cases. In other words, satisfaction of Equations 12 or 14 is a necessary but not sufficient reason to accept the calculated value of the intergreen interval requirement at a given intersection. The "solution" to Equation 12 or 14 can at best be viewed as an initial estimate of the intergreen interval requirement, subject to adjustments based on the resulting speed-location diagram implied by this initial estimate.

IMPORTANCE OF SPEED-LOCATION FRAMEWORK: AN ILLUSTRATION

The fundamental importance of the speed-location framework to guide the interpretation of experimental results is illustrated by using the data reported by Stimpson et al. (*6*). Their research attempted to determine whether changing the time duration of yellow signals (referred to as change interval) should affect the frequency of potential conflicts. They defined a potential conflict to exist whenever "the last-to-cross vehicle spent at least 0.2 seconds in the intersection past red onset." To accomplish their objective, they selected two suburban intersections, one in Bethesda, Maryland, the other in Atlanta, Georgia. They carefully selected the experimental sites to ensure certain conditions, including average approach speed near 30 mph, "short" yellow durations (less than 5 sec), reasonably isolated intersections with pretimed signals, four-legged intersections with negligible grade, and good pavement surfaces. The existing yellow durations were 4.7 and 4.3 sec for the Maryland and the Georgia intersections, respectively. For each intersection, they compared the percentage of potential intersection conflicts when the existing yellow was present against the percentage observed when the yellow signal was extended. The percentage of potential conflicts was defined as the ratio of last-to-cross "decision vehicles" (see below) that spent at least 0.2 sec in the intersection past red onset to the total number of decision vehicles that were last to cross.

The yellow at the Maryland location was extended from 4.7 to 6.0 sec, which was the maximum duration acceptable to the responsible traffic engineer. In an attempt to ensure comparability of the results obtained at the two sites, they extended the yellow at the Georgia intersection to "produce a percentage increase of similar magnitude" (i.e., from about 4.3 to about 5.6 sec, with minor variations). At each intersection, before-and-after data were collected separately for peak and off-peak conditions on dry pavements. Observations relating to wet pavement conditions were also collected at the Maryland site. The data were collected via lapse photography and corresponded to vehicles that occupied a "catch zone" 2 sec prior to the onset of yellow. The catch zone was selected so that it "included the dilemma zone at most approach speeds" on the following basis (*6*):

> The upstream extremity of the catch zone was chosen as the point from which a car with an initial speed of 10 mph in excess of the local average speed could come to a full stop at the traffic signal using an average deceleration of 0.25 g (8 ft/sec^2). The downstream extremity was chosen at a point from which a vehicle traveling 10 mph below the average at yellow onset could just clear the cross street prior to red onset. At the

> Maryland site the catch zone extended from 65 feet to 320 feet and at the Georgia site from 25 feet to 320 feet.

The recording and data reduction procedures were summarized as follows (*6*):

> Filming commenced at least two seconds prior to yellow signal onset and continued until all vehicles initially in the catch zone either stopped or cleared the intersection. . . . For the purpose of this study, a vehicle was called a decision vehicle if, in a particular approach lane, it was 1) the first vehicle to stop, or 2) the last vehicle to cross the intersection. Data collection continued until about 150 decision vehicles were obtained at each site under each experimental condition.

A reduction in the percentage of potential conflicts was observed when the yellow signal was extended at each of the two locations and for all experimental conditions investigated. The results corresponding to peak and off-peak conditions on dry pavements are shown in Table 1. Of special interest here is the comparison of the results between the two sites (*6*):

> The results . . . show that potential conflict percentages differed between the two sites both with the initial and extended yellow durations. These differences undoubtedly reflect differences between the two sites in terms of geometry, approach speed and traffic volume . . . but there is not at present quantitative relationships that would predict potential conflict frequency in terms of these, and possible other, factors.

We agree with Stimpson et al. as to the possible factors that gave rise to the observed differences between the two sites. We contend, however, that the speed-location diagrams reflecting the two yellow durations at the two intersections studied could contribute to an explanation of their findings. Such speed-location diagrams for the Maryland and Georgia sites are shown in Figures 5 and 6. The X_0 curves shown were based on a perception-reaction time of 1 sec, a deceleration rate of 10 ft/sec^2, and vehicle length of 20 ft. Also shown on each figure are the "catch zones" within which experimental data were collected as described. It was not possible to show similar ranges of observed speeds because only the average approach speeds are given by Stimpson et al.; these averages are included in the two figures.

Even a cursory inspection of the two figures is sufficient to pinpoint the prevailing differences at the two intersections and to provide a reasonable explanation of the experimental findings: Given the initial yellow durations at the two inter-

TABLE 1 SUMMARY OF RESULTS (*6*)

a) Maryland Intersection

	Before (T = 4.7 sec.)	After (T = 6 sec.)
Peak	19%	2%
Offpeak	15%	1%

b) Georgia Intersection

	Before (T = 4.4 sec.)	After (T = 5.6 sec.)	After (T = 5.8 sec.)
Peak	90%	21%	(*)
Offpeak	63%	(*)	19%

(*) Not Applicable

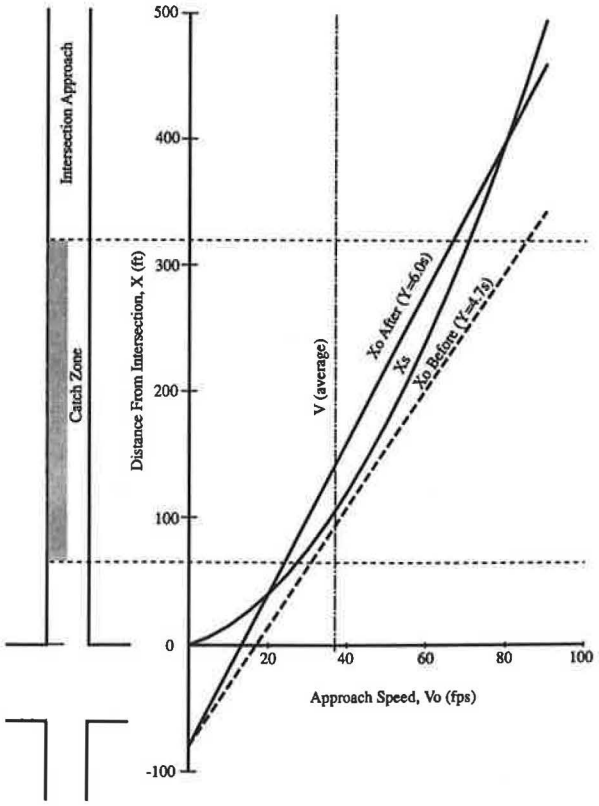

FIGURE 5 Speed-location diagram of Maryland site (*6*).

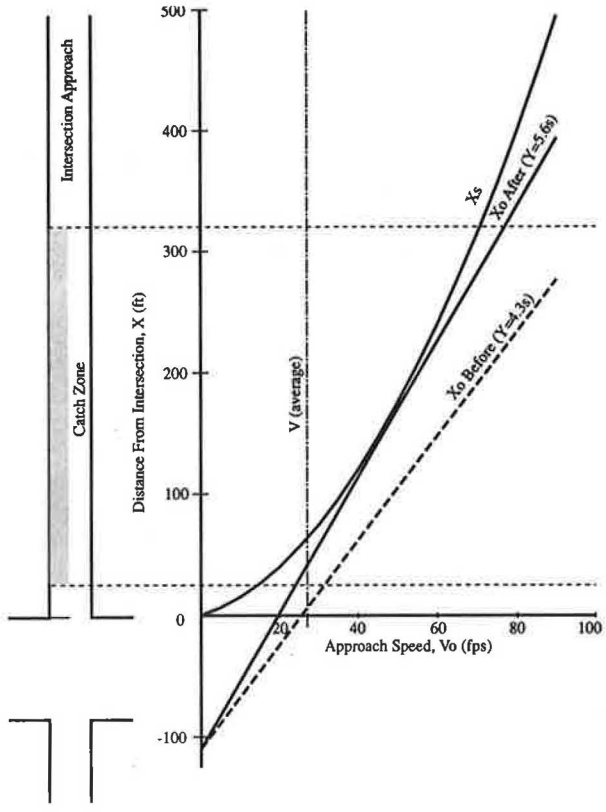

FIGURE 6 Speed-location diagram of Georgia site (*6*).

sections, it is eminently obvious that dilemma zones of significant lengths existed at the Georgia site along the entire length of the catch zone. This intersection's 90-ft width (as reflected by the X_0-intercept) plays a significant role in the presence of dilemma zones. By contrast, at the Maryland site, dilemma zones of any significant length (about 20 ft or more) appear only on the high-speed end of the graph. Thus, the two simple diagrams are sufficient in this case to present the traffic engineer with a clear warning that the Georgia site was likely to experience a higher percentage of potential conflicts than the Maryland site.

The two diagrams are also consistent with the experimental results obtained after the extension of the yellow durations. In the Maryland case, no dilemma region remained over the length of the catch zone after resetting the yellow light; this is consistent with the finding that the percentage of potential conflicts dropped to essentially zero. By contrast, the diagram corresponding to the Georgia site after the signal extension shows that dilemma regions were present at both ends of the catch zone. This finding is consistent with the observed persistence of vehicles in potential conflict at the Georgia site and not at the Maryland site.

Despite the meticulousness with which the researchers attempted to establish their experimental design, they inadvertently failed to account for significant differences at the two intersections that could have been readily disclosed through the use of speed-location diagrams. In other words, in this case, ensuring an equivalent extension of the two signals was clearly not sufficient to render the two intersections comparable. We have discovered similar inadvertent flaws in several experimental designs reported in the literature.

ADDITIONAL COMMENTS

To some degree, the percentages of potential conflicts reported by Stimpson et al. were dependent on the specific choice of catch zones to which their observations were confined. The Maryland site catch zone practically excluded the dilemma region adjacent to the stop line which arises at low approach speeds; the Georgia site catch zone included part of that region. Without a doubt, the observed percentages would have been different had the two catch zones been extended to the stop-line location, including among the decision vehicles those facing the near-side dilemma zones. The observed percentages, of course, would also depend on the actual distribution of observed decision vehicles on the speed-location plane, information that is currently unavailable.

Experimental data can be presented in a convenient and meaningful way by showing, on speed-location diagrams such as those in Figures 5 and 6, the initial speeds and positions of the observed vehicles. Distinct symbols may be used to associate to each observed vehicle its subsequent action (e.g., whether it stopped, whether it cleared the intersection during yellow, or whether it cleared the intersection on red). At a minimum, such practice is capable of transmitting needed information about the intersection under study and its signal characteristics, the sample of vehicles used, and the specific actions taken by the observed drivers. The X_0 and X_S curves can provide guidance in assessing the actions of observed drivers, for example, whether a driver that chose to clear the intersection could have stopped comfortably. Most authors report the duration of the intergreen intervals prevailing during their experimental sessions but typically fail to provide sufficient information to even discern whether dilemma and/or option regions were present at the sites investigated. Moreover, information relating to the distribution of vehicles in the experimental samples used between the various regions of the speed-location diagram is typically absent. Sample selection procedures so differ between researchers that their conclusions are often impossible to compare. For example, Olson and Rothery (*18*) recorded free-flowing vehicles over catch zones not extending to the stop line and consciously disregarded vehicles traveling at considerably lower speeds. As described above, Stimpson et al. (*6*) considered all vehicles occupying similar catch zones at the onset of yellow. Williams (*10*) reports that 816 close-decision vehicles were recorded at a single intersection but gives no further description of his methods; Chang et al. (*20*) sampled vehicles approaching faster than 20 mph. Moreover, with the possible exception of May (*1,2*), who has reported his observed sample on speed-location diagrams tailored to the specific characteristics of the intersections studied, researchers tend to merely report some statistical descriptor of their sample speeds along with the signalization existing during their experimental sections. Typically, sufficient information to plot the corresponding speed-location diagrams is unavailable in research reports. Sampling inconsistencies are probably a main source of the conflicting conclusions reported by various researchers, particularly those who attempt to generalize data applicable to restricted scopes and those who attempt to discover behaviorally sound models through blind regression analyses using data from incompatible sources. Without the kind of site-specific information that can be depicted by speed-location diagrams, the conclusions drawn by researchers must remain suspect, particularly when, as in the case of the preceding illustration, the researchers attribute nonconforming findings to unknown factors.

FURTHER REFINEMENTS

Further refinements to the basic speed-location diagram described here are possible. For example, an extended diagram can aid in the depiction of the implications of legal requirements, particularly those that are implied by the permissive yellow rules. It can also clarify the role and implications of competing proposals regarding the division of the intergreen interval into yellow and all-red components, including the current ITE-proposed recommended practice (*4,12*). However, length limitations preclude a full discussion of these important questions here.

CONCLUSIONS

This paper traces the chronological development of the most commonly used intergreen interval formula and identifies the major differences of interpretation that are prevalent among traffic engineers and researchers. By far, the most critical difference lies in the disparity between the interpretation presumed by ITE and that originally proposed by Gazis et al. Although both interpretations are based on the same equa-

tions of motion, the former strongly implies that, given appropriate design parameters, the mere solution of the design equation is sufficient to yield an intergreen interval duration which eliminates the dilemma zone problem. This paper clearly shows that this is not the case and that, properly used, the speed-location framework can be an invaluable tool that is capable of enhancing the traffic engineer's judgment and evaluation of initial estimates of intergreen timings at specific intersections. It can also provide a consistent means of reporting research-related data in a manner that can aid the interpretation, understanding, and comparison of formerly incompatible research findings. It is therefore strongly recommended that the speed-location diagram be adopted as a standard tool by traffic engineering practitioners and researchers. Given existing computer technology and graphics software, incorporating the speed-location diagram in practice would be relatively easy.

ACKNOWLEDGMENT

The authors wish to thank Kay Kasamoto for her editorial help.

REFERENCES

1. A. D. May. *Study of Clearance Interval at Traffic Signals*. Institute of Transportation and Traffic Engineering, University of California, Berkeley, 1967.
2. A. D. May. Clearance Intervals at Traffic Signals. In *Highway Research Record 221*, HRB, National Research Council, Washington, D.C., 1968, pp. 47–71.
3. T. M. Matson. The Principles of Traffic Signal Timing. *Trans., 18th Annual Safety Congress*, National Safety Council, Vol. 3, 1929, pp. 109–139.
4. ITE Technical Committee 4A-16. Proposed Recommended Practice—Determining Vehicle Signal Change Intervals: Part I. *ITE Journal*, July 1989, pp. 29–32.
5. D. Gazis, R. Herman, and A. Maradudin. The Problem of the Amber Signal in Traffic Flow. *Operations Research*, Vol. 8, No. 1, Jan.–Feb. 1960.
6. W. A. Stimpson, P. A. Zador, and P. J. Tarnoff. The Influence of Time Duration of Yellow Traffic Signals on Driver Response. *ITE Journal*, Nov. 1980, pp. 22–29.
7. F. R. Hulscher. The Problem of Stopping Drivers After the Termination of the Green Signal at Traffic Lights. *Traffic Engineering and Control*, Vol. 25, No. 3, March 1984, pp. 110–116.
8. H. K. Evans (ed.). *Traffic Engineering Handbook*. ITE, New Haven, Conn., 1950.
9. P. S. Parsonson and A. Santiago. Traffic Signal Change Interval Must Be Improved. *Public Works*, Sept. 1981, pp. 110–113.
10. W. L. Williams. Driver Behavior During the Yellow Interval. In *Transportation Research Record 644*, TRB, National Research Council, Washington, D.C., 1977, pp. 75–78.
11. T. M. Matson, W. S. Smith, and F. W. Hurd. *Traffic Engineering*. McGraw-Hill Book Company, Inc., New York, Toronto, London, 1955.
12. ITE Technical Committee 4A-16. Proposed Recommended Practice—Determining Vehicle Signal Change Intervals: Part II, Literature Review and Committee Deliberations. ITE, Washington, D.C., c. 1989.
13. H. H. Bissell and D. L. Warren. The Yellow Signal is not a Clearance Interval. *ITE Journal*, Feb. 1981, pp. 14–17.
14. J. E. Baerwald (ed.). *Traffic Engineering Handbook*. 3rd ed. ITE, Washington, D.C., 1965.
15. R. H. Wortman and T. C. Fox. A Reassessment of the Traffic Signal Change Interval. In *Transportation Research Record 1069*, TRB, National Research Council, Washington, D.C., 1986, pp. 62–68.
16. F. B. Lin. Timing Design of Signal Change Intervals. In *Transportation Research Record 1069*, TRB, National Research Council, Washington, D.C., 1986, pp. 46–51.
17. J. E. Baerwald. *Transportation and Traffic Engineering Handbook*. ITE, New Jersey, 1976.
18. P. L. Olson and R. Rothery. Driver Response to Amber Phase of Traffic Signals. In *Highway Research Bulletin 330*, HRB, National Research Council, Washington, D.C., 1962, pp. 40–51.
19. W. S. Homburger. *Transportation and Traffic Engineering Handbook*. ITE, New Jersey, 1982.
20. M. S. Chang, C. J. Messer, and A. J. Santiago. Timing Traffic Signal Change Intervals Based on Driver Behavior. In *Transportation Research Record 1027*, TRB, National Research Council, Washington, D.C., 1985, pp. 20–30.

DISCUSSION

FENG-BOR LIN
Civil and Environmental Engineering, Clarkson University, Potsdam, N.Y. 13699-5710

The authors should be commended for their attempt to resolve the controversy surrounding the timing of the intergreen interval. Their paper is based on the argument that the dilemma zone concept as advanced by Gazis et al. (*1*) must always be used in evaluating a given intergreen interval. This bias has limited the scope of their literature review and discussions.

Much of the controversy in timing the intergreen interval stems from several equations suggested by ITE over the years. In 1985, for example, ITE proposed that the following equation be used to determine the yellow interval (*2*):

$$Y = t + V/[2\ (a + gG)] \tag{1}$$

where

Y = length of yellow interval,
t = driver perception/reaction time,
V = vehicle approach speed,
a = deceleration rate,
G = grade of approach lane, and
g = gravitational acceleration.

Although the root of this equation can be traced back to the dilemma zone concept, ITE (*2*) has also indicated that the primary measure of effectiveness for the yellow interval is the percentage of vehicles entering the intersection after the yellow interval expires. This measure of effectiveness reflects the need to reduce the potential of right-angle collisions. In other words, the real intention of Equation 1 is to ensure that the yellow interval is long enough to allow most drivers who are faced with a yellow light to come to a stop rather than continue to enter the intersection after the red onset. Unfortunately, because they are derived from dilemma considerations, Equation 1 and other similar equations are not compatible with this intended timing requirement. This incompatibility is one reason for the different interpretations of such equations; it is also a weakness of ITE's equations. Several studies of driver behavior (*3–5*) have consistently shown that the yellow interval needed to prevent a high per-

centage of drivers from entering on red is independent of vehicle approach speed. The underlying reason for this phenomenon is probably drivers' willingness to tolerate (or apply) greater deceleration rates at higher approach speeds. In any case, the findings of these studies have prompted suggestions to use a uniform yellow interval. The authors have ignored this side of the controversy.

Because the yellow interval needed to prevent most drivers from entering on red is independent of vehicle approach speed, the consideration of dilemma situations is really not that important. For the same reason, the use of the speed-location diagram as illustrated by the authors becomes an unnecessary exercise. For example, the experimental results reported in Table 1 can be logically explained in terms of yellow interval demand, which refers to the length of the yellow interval that is needed in a change interval to prevent drivers from entering on red. A typical cumulative distribution of yellow interval demand is shown in Figure 7. This distribution is based on observed driver behavior at intersections that are controlled with pretimed signals (*5*). It should be noted that such a distribution is independent of vehicle approach speed. Assuming that the drivers at the Maryland and the Georgia sites exhibited the same behavior as that shown in Figure 7, this figure can be used to predict the impact of the yellow interval on the conflict potential at these sites. For example, the Maryland site has a clearance distance of about 75 ft. If vehicles moved across this intersection at an average speed of 30 mph, then it would take an average of 1.7 sec to clear the intersection. With a 4.7-sec yellow interval during off-peak hours, this clearance time requires a driver to enter the intersection within the first 3.2 sec (e.g., 4.7 − 1.7 + 0.2) of the yellow interval in order to avoid occupying the intersection for more than 0.2 sec after the red onset. Figure 7 shows that in 55 percent of the change intervals, drivers will continue to enter the intersection after 3.2 sec of the yellow interval has elapsed. In other words, late entries can be expected to exist in 45 percent of the change intervals. Similar estimates can be obtained for the various conditions reported by the authors in Table 1. These estimates and their relationships to the respective conflict potentials are shown in Figure 8. Due to the lack of actual data on clearance time, the estimated percent-

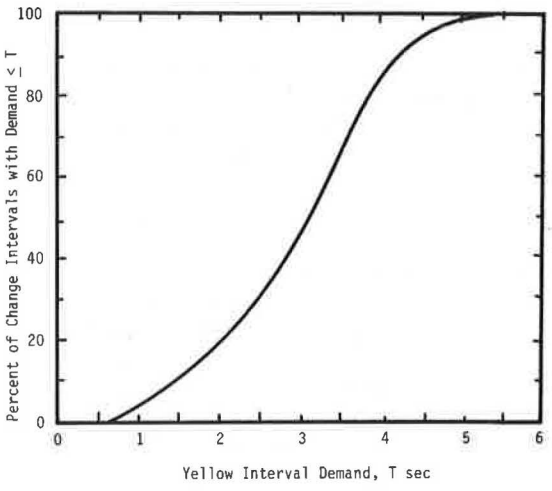

FIGURE 7 Cumulative distribution of yellow interval demand.

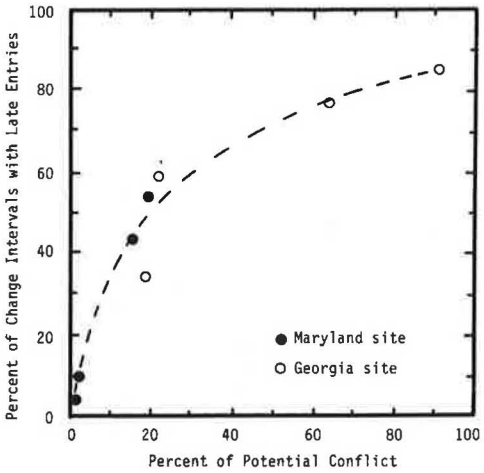

FIGURE 8 Relationship between conflict potential and percentage of change intervals with late entries.

age of change intervals with late entries are based on a clearance speed of 30 mph for off-peak hours and 25 mph for peak hours. This figure clearly shows that the impact of changing yellow interval can be reasonably predicted without knowing whether there is a dilemma.

In light of the predominant concern about right-angle collisions, I believe that the dilemma zone concept is elusive and its importance overstated. A situation that is a dilemma to one driver is not necessarily a problem to another.

REFERENCES

1. D. Grazis, R. Herman, and A. Maradudin. The Problem of the Amber Light in Traffic Flow. *Operations Research*, Vol. 8, No. 1, Jan.–Feb. 1960.
2. ITE Technical Committee 4A-16. Proposed Recommended Practice: Determining Vehicle Changed Intervals. *ITE Journal*, May 1985, pp. 61–64.
3. M. S. Chang, C. J. Messer, and A. J. Santiago. Timing Traffic Signal Change Intervals Based on Driver Behavior. In *Transportation Research Record 1027*, TRB, National Research Council, Washington, D.C., 1985, pp. 20–30.
4. R. H. Wortman and T. C. Fox. A Reassessment of the Traffic Signal Change Interval. In *Transportation Research Record 1069*, TRB, National Research Council, Washington, D.C., 1986, pp. 62–68.
5. F. B. Lin and S. Vijaykumar. Timing Design of Signal Change Interval. *Traffic Engineering and Control*, Vol. 29, No. 10, Oct. 1988, pp. 531–536.

AUTHORS' CLOSURE

We thank Lin for his discussion and the opportunity to respond to his comments. Contrary to his expectation, our paper did not attempt to resolve the controversy but rather to propose a common framework that would help avoid the prevalent problem we discovered in our literature search relating to differences in definitions, methods of measurement, sampling, and so on. Among the advantages of adopting the speed-location diagram as the basic framework is that it can explicitly present aspects of the problem that have been persistently

considered by almost all, if not all, researchers in one way or another. These aspects include the speed and location of observed vehicles at the onset of the intergreen interval, the width of the intersection, the actions taken by observed drivers and their consequences, and the equations of motion governing the stop-or-go decision. Other methods of presenting observed data either omit some of these factors or consider them implicitly.

It is unfortunate that Lin failed to make a distinction between ITE's very restricted interpretation of the problem and the larger framework afforded by the full speed-location diagram, and that he dismissed the usefulness of the latter based on the shortcomings of the former, a practice against which we clearly warned. Of particular concern is the fact that the ITE formula is based on a design situation that involves a vehicle that, at the onset of the intergreen interval, happens to be traveling at the selected design speed and happens to be located behind the stop line at a distance exactly equal to the stopping distance corresponding to that speed. ITE's assumption, which is evident in Lin's Equation 1, is that if the signal were to be timed for this design situation, the conditions faced by vehicles approaching at other speed and location combinations would be covered. Our paper showed that this is not the case and that the speed-location diagram can provide valuable guidelines which can aid the timing engineer's judgment. More definite timing guidelines must await further study of experimental data in the context of the speed-location diagram.

We have also shown that the presentation of experimental data on the speed-location diagram can help researchers (particularly timing engineers) to systematically interpret their results by showing the intersection width, the characteristics of the signal, the speed and location of the subject vehicles at the onset of yellow, the drivers' decision to stop or go, and the consequences of these decisions. Our review has revealed that a large part of the conflicting conclusions reported in the literature can be traced to differences in sampling which can be made explicit through the use of the speed-location diagram. This point is discussed in the "Additional Comments" section of our paper, which we urge Lin to study more carefully.

Lin criticizes our paper for not discussing the potential effectiveness of adopting a constant yellow interval. Even though our work in this area has included the question of the division of the intergreen interval into yellow and all-red and other important issues, it was not possible to present our findings on these matters in a single paper because of length limitations. Our failure to emphasize this point may have contributed to Lin's primary concern with the yellow interval. A constant yellow would impart a degree of certainty to drivers who are not familiar with the intersection; what may be considered to be a reasonable action by a familiar driver may not always be expected to be so in the case of the unfamiliar driver who has no idea as to the timing or operational characteristics of an intersection visited for the first time. Whether a constant yellow, however, would cause other difficulties, such as excessive all-red intervals at some intersections, requires additional research.

Lin argues that ITE's proposal of using his Equation 1 to calculate the required yellow is inappropriate because several studies have "shown that the yellow interval needed to prevent a high percentage of drivers from entering on red is independent of vehicle approach speed." We agree that the ITE proposal is inappropriate, but our objection to it lies in the fact that the ITE formula is based on using a single approach speed rather than on examining the resulting conditions over the full range of speed and location conditions. As explained later, Lin's model also suffers from this shortcoming.

The studies cited by Lin prefer the use of time to reach the stop line as a superior criterion for the timing of the yellow interval. As usually defined, this time depends on both the vehicle's speed and its location at the onset of yellow, rather than speed alone. The ability of the speed-location diagram to show both vehicle speed and location at the onset of yellow is precisely one of the reasons that we propose its adoption. Moreover, as usually defined, the time to reach the stop line can be explicitly shown as the slope of a straight line drawn from the origin of the speed-location diagram to the point representing the initial conditions corresponding to a particular vehicle in the sample. Alternatively, a straight line from the origin with a particular slope would divide the speed-location space into two regions representing, respectively, the conditions that allow vehicles to enter the intersection within the time represented by the slope and those that do not. Bissel and Warren (*1*), in fact, used this concept with a slope that was equal to the duration of yellow. Our studies have shown that superposition of such a line on the speed-location diagram described in the paper can further enhance its usefulness. We have termed a diagram that includes this line and a representation of the division of the intergreen interval into yellow and all-red (see below) as the "expanded speed-location diagram," which we intend to describe more fully in a subsequent paper.

Figure 9 shows such an expanded speed-location diagram for the conditions prevailing at one of the sites studied by Lin (*2*). The dashed line has a slope that is equal to the duration of the existing yellow. Vehicles whose speed and location at the onset of yellow place them above this line cannot reach the stop line (i.e., cannot enter the intersection) at constant speed before expiration of the yellow interval, whereas vehicles that plot below the dashed line can. The diagram also shows two clearing lines: one based on the existing intergreen interval consisting of the sum of the yellow and all-red subdivisions (T = 5.9 sec), the other on the yellow portion only (Y = 3.0 sec). The mean approach speed was reported to be 30.6 mph with a 15th percentile of 24.2 mph and a 95th percentile of 35.8, but no information is provided regarding how speeds were measured, the particular catch zone used, or the distribution of the vehicles in the sample on the diagram. Of 11 sites studied, this intersection approach was observed to have the longest maximum change interval requirement of 8.9 sec, that is, at least one observed vehicle took this long from the onset of yellow to clear the intersection. Figure 9 shows that severe dilemma regions existed at this intersection approach; it is also noteworthy that the existing timing was shorter than that recommended by ITE. In other words, the speed-location diagram can be drawn irrespective of the method used in setting the timing and it should not be viewed as wed to the ITE method as Lin implied.

Had Lin chosen to report his sample of vehicles on the speed-location diagram, a more systematic investigation of

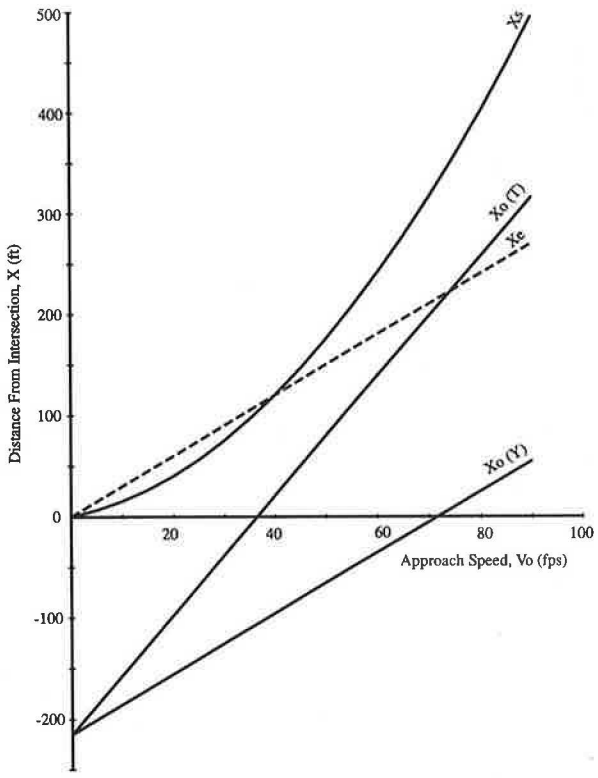

FIGURE 9 Expanded speed-location diagram for site studied by Lin.

the observed drivers could be possible. Instead, he chose to combine the data obtained at the 11 sites with widely varying widths and signal characteristics in order to examine "driver's aggregate needs" and to perform regression analyses in the hope of discerning a universally applicable relationship between the change interval requirement and the average approach speed. A year later, Lin et al. (*3*) used a similar procedure with additional data to arrive at a different change interval design equation. They also proposed the use of a yellow interval requirement distribution, similar to that shown by Figure 7 in Lin's discussion, to be used in timing the duration of yellow. This concept was further discussed by Lin and Vijaykumar (*4*). An examination of the yellow interval requirement distributions reported in these two papers (*3*,*4*) for several of the sites studied by Lin and his co-workers reveals a great variability between them. For example, the maximum observed yellow interval requirement (shown as approximately equal to 5.5 sec in Lin's discussion) actually ranged from a low of 3.3 sec at one site to a high of 6.4 sec at another (*4*). Thus Lin's claim that such a typical distribution exists cannot be substantiated at this time. In addition, in his illustration on how to apply his model, he calculated and subtracted from the intergreen interval a crossing time based on a single assumed clearing speed. This is reminiscent of the ITE interpretation of crossing time based on a single design speed, which is inappropriate. Moreover, the significant discrepancies evident on his Figure 8 between his calculated "percent of change intervals with late entries" and the "percent of potential conflicts" observed by Stimpson et al. (*5*) provide sufficient reason to question the validity of Lin's typical distribution of yellow interval requirements.

Finally, it is instructive to examine an explanation offered by Lin and Vijaykumar (*4*) of their finding that their observed maximum change interval demands were, on average, 53 percent longer than the actual settings of the change interval. They stated that "the drivers either could not or would not comply with the traffic regulation that requires one to clear the intersection by the time the change interval expired." In this connection, a paper cited in Lin's discussion observed that "the presence of a police vehicle at the site significantly reduced the percentage of vehicles entering on red" (*6*). In other words, Lin's preferred model cannot distinguish between observed noncomplying actions that are under the driver's control from those that are not. The speed-location representation can be more helpful in this respect.

In conclusion, consideration of Lin's discussion strengthens the case for adopting the speed-location diagram as a tool for evaluating research results. Lin and his co-workers are among the few researchers who have actually conducted field experiments. For this they should be commended. We encourage them, however, to report their data in the context of the speed-location framework so that progress can be made toward the resolution of the controversy.

REFERENCES

1. H. H. Bissell and D. L. Warren. The Yellow Signal Is Not a Clearance Interval. *ITE Journal*, Feb. 1981, pp. 14–17.
2. F. B. Lin. Timing Design of Signal Change Intervals. In *Transportation Research Record 1069*, TRB, National Research Council, Washington, D.C., 1986, pp. 46–51.
3. F. B. Lin, D. Cooke, and S. Vijaykumar. Utilization and Timing of Signal Clearance Interval. In *Transportation Research Record 1114*, TRB, National Research Council, Washington, D.C., 1987, pp. 86–95.
4. F. B. Lin and S. Vijaykumar. The Timing Design of Signal Change Interval. *Traffic Engineering and Control*, Vol. 29, No. 10, 1988, pp. 531–536.
5. W. A. Stimpson, P. A. Zador, and P. J. Tarnoff. The Problem of the Amber Signal in Traffic Flow. *Operations Research*, Vol. 8, No. 1, Jan.–Feb. 1960.
6. R. H. Wortman and T. C. Fox. A Reassessment of the Traffic Signal Change Interval. In *Transportation Research Record 1069*, TRB, National Research Council, Washington, D.C., 1986, pp. 62–68.

The authors are solely responsible for the contents of this paper.

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