Field Studies of Temporary Pavement Markings at Overlay Project Work Zones on Two-Lane, Two-Way Rural Highways

CONRAD L. DUDEK, R. DALE HUCHINGSON, F. THOMAS CREASEY, AND OLGA PENDLETON

In response to FHWA's proposed rule requiring that all states use 4-ft pavement markings on 40-ft centers as temporary markings in highway work zones, NCHRP awarded a research contract to the Texas Transportation Institute to conduct field studies to compare the safety and operational effectiveness of 1-ft, 2-ft, and 4-ft temporary broken line pavement markings in work zones. The following scope and test conditions were specified by NCHRP: (a) surfacing operations on two-lane, two-way facilities; (b) field sites involving pavement overlays (not seal coats); (c) data collection during hours of darkness; (d) dry roadway conditions; (e) sites with both tangent and curve sections; (f) centerline stripe only (no edgelines); (g) use of a 40-ft pavement marking cycle; and (h) field tests in real or staged work zones that are open to traffic. Field studies were conducted at night at seven pavement overlay project sites on two-lane, two-way rural highways in Arkansas, Colorado, Oklahoma, and Texas. Traffic stream measures of effectiveness included vehicle speeds, lateral distance from the centerline, lane straddling, and erratic maneuvers. In-vehicle studies using paid driver subjects were conducted to supplement the traffic stream evaluation. The 1-ft and 2-ft striping patterns on 40-ft centers performed as well as the 4-ft pattern for centerline striping at night for the conditions studied: pavement overlay projects on rural two-lane, two-way highways with 2.0 degree horizontal curvature, level to rolling terrain, and average speeds between 50 and 62 mph. Although the driver subjects at six sites rated the 1-ft pattern to be the least effective on the average, there was no statistical difference in mean ratings or rankings among the three patterns.

The cost of temporary traffic control is significant for many construction, maintenance, and utility projects. With the prospects of continued inflation, limited resources, and high interest rates, it is imperative that all aspects of temporary traffic control be evaluated for economy in application and benefits to the public.

FHWA has issued guidelines and proposed changes to the *Manual on Uniform Traffic Control Devices* (MUTCD) regarding temporary markings for construction and maintenance areas (*1*). Markings that are less than the full standard marking pattern (10-ft stripe on 40-ft centers) would be permitted for broken lines, but the proposed changes would require a minimum pattern of 4-ft stripes on 40-ft centers (36-ft gaps), which is more than double what many states now specify. There has been concern that if the 4-ft markings were adopted as the national standard, they would significantly increase project costs. Table 1 is a summary of data abstracted from a survey conducted by the Traffic Engineering Section of the Arizona Department of Transportation in 1986. The number of states using each of 15 different temporary pavement striping patterns is presented. NCHRP awarded a contract to the Texas Transportation Institute (TTI) to determine whether the proposed 4-ft markings on 40-ft centers would actually result in significant safety and operational improvements in comparison to current practice (2).

The Texas Transportation Institute, Texas A&M University, College Station, Tex. 77843.

The specific objective of the research was to conduct field studies comparing the safety and operational effectiveness of 1-ft, 2-ft, and 4-ft temporary broken line pavement markings in work zones. To ensure that the findings would be applicable to situations in which this type of marking is most typically used, the following scope and test conditions were identified by NCHRP:

- Surfacing operations on two-lane, two-way facilities;
- Field sites involving pavement overlays (not seal coats);
- Data collection during hours of darkness;
- Dry roadway conditions;
- Sites with both tangent and curve sections;
- Centerline stripe only (no edgelines);
- Use of a 40-ft pavement marking cycle; and
- Field tests in real or staged work zones that are open to traffic.

STATE OF THE ART

A review of the literature revealed a variety of research projects on work zone traffic control. However, little information was available on the relative effectiveness of temporary pavement marking patterns in work zones.

Godthelp and Riemersma used a theoretical analysis to estimate the effectiveness of particular delineation systems as a reference in perceiving course and speed (*3*). Although this work is very general, it does provide insight into the interactions of the driver and the driving environment. The authors

TABLE 1 SUMMARY OF TEMPORARY PAVEMENT MARKING PATTERN PRACTICE, 1986

Length of Stripe (ft)	Length of Gap (ft)	Striping Interval (ft)	Number of States
10	30	40	11[a]
8	32	40	1
5	95	100	1
4	36	40	8
3	37	40	1
3	77	80	1
2	18	20	1
2	38	40	7
2	48	50	6
2	78	80	1
2	98	100	1
1	24	25	2
1	39	40	6
1	74	75	1
1	79	80	1
States using separate markings for curves			7
States using temporary edgelines			26

NOTE: Survey conducted by the Arizona Department of Transportation, Traffic Engineering Section (*2*).
[a]Five of the 11 states allow stripes less than 4 ft long under specified conditions.

TABLE 2 TEMPORARY PAVEMENT MARKING TREATMENTS EVALUATED BY TTI IN PROVING GROUNDS SETTING

Treatment	Description
1[a]	4-ft stripes (4 in. wide) with 36-ft gaps (control condition)
2[a]	2-ft stripes (4 in. wide) with 38-ft gaps
3[a]	8-ft stripes (4 in. wide) with 32-ft gaps
4[a]	2-ft stripes (4 in. wide) with 18-ft gaps
5[a]	Four nonreflective RPMs at 3¹/₃-ft intervals with 30-ft gaps *and* reflective marker centered in alternate gaps at 80-ft intervals
6[a]	Three nonreflective and one reflective RPMs at 3¹/₃-ft intervals with 30-ft gaps
7	2-ft stripes (4 in. wide) with 48-ft gaps
8	Treatment 2 *plus* RPMs at 80-ft intervals
9[a]	Two nonreflective RPMs at 4-ft intervals with 36-ft gaps *plus* one reflective RPM centered in each 36-ft gap
10	1-ft stripes (4 in. wide) with 19-ft gaps

[a]Treatments evaluated both day and night.

point out the obvious fact that work zones represent discontinuities for drivers in terms of driving speed and roadway characteristics and consequently place special demands on the traffic control devices used in these areas. Godthelp and Riemersma also conducted laboratory experiments to preview the guidance effectiveness of delineation devices (*4*). Their findings suggested that placement of delineators at a level lower than the driver's eye height improved delineation efficiency and that chevron panels were particularly effective if other devices tended to be somewhat haphazardly placed.

Raised pavement markers (RPMs) for construction zone delineation were examined in Arkansas by Spencer (*5*). This study reported that RPMs provided excellent wet weather and nighttime reflectivity and appeared to be an effective means of maintaining safe traffic flow in work zones. Niessner (*6*) reviewed the practices of nine state highway agencies concerning the use of RPMs for temporary delineation in work zones. A wide variety of projects were included. The nine state highway agencies reported that the RPMs provided excellent nighttime temporary delineation, particularly on wet roads. In addition, the delineation was low cost and required little or no maintenance. In two projects reported by Niessner, an accident reduction occurred. Officials in the majority of the states said that they would continue to use the RPMs in construction zones after the study had been concluded.

The Texas Transportation Institute (TTI) investigated candidate temporary pavement marking treatments for use at work zones (Table 2) by determining the effects of each on various measures of driving performance (*7*). The studies were conducted on a 6-mi test track at TTI's proving ground facility. Ten candidate temporary pavement marking treatments were evaluated during daylight hours, and seven of the ten candidates were also evaluated at night. The candidate treatments included patterns with stripes, RPMs, and combinations of both. Treatment 1 (4-ft stripes with 36-ft gaps) was considered to be the control condition in the studies.

The major findings from the daytime proving ground studies were as follows:

- The vehicle speed and distance data failed to provide any basis for selection among the 10 treatment conditions. Because of the large variability within subjects and the small magnitude of change in the measures of effectiveness (MOEs), the analysis of the objective data failed to reveal any practical significant difference in treatments.
- Two treatments with short (2-ft) stripes and long gaps (48- and 38-ft intervals) were associated with missed curves and with a few wide deviations to the right of the centerline.
- The subjective ratings tended to support the data just mentioned. Drivers indicated that it was difficult to follow curves with short stripes or long gaps and preferred the 8-ft stripe with 32-ft gap pattern and the RPMs.

The major findings and conclusions of the nighttime studies were as follows:

- Speed and distance performance data for the nighttime studies were not sufficiently different to provide a basis for ranking the treatments. Speed profiles for night driving were comparable to those for the daytime studies.
- Erratic maneuver data also revealed no significant differences with respect to treatments.
- Drivers rated the 8-ft stripes with 32-ft gaps as the best and the 2-ft stripes with 38-ft gaps as the poorest of the four striping patterns tested. The three RPM treatments tested were all judged by drivers to be highly effective.
- Drivers rated the baseline treatment (4-ft stripes with 36-ft gaps) to be inferior to the three RPM treatments tested.
- In general, the nighttime studies supported the findings of the daytime studies in ratings of effectiveness. However, neither study found that the performance data provided any basis for ranking the seven treatments.

The TTI researchers emphasized that studies performed on proving grounds are no substitute for real-life field studies. Proving ground studies can help identify and eliminate candidate treatments that are considerably ineffective relative to the others. However, because subject drivers tend to do their best when tested in a proving ground setting, the test is not generally sensitive enough to discern small differences between candidate treatments. Field studies must be conducted to measure these differences.

FIELD STUDIES

Field Study Plan

A brief description of the study plan is given in the following sections. More complete details are provided elsewhere (2).

Study Sites

Field studies were conducted at seven pavement overlay construction projects on two-lane, two-way rural highways. The allocation of study sites was four in Texas and one site each in Arkansas, Colorado, and Oklahoma. The order of studies was as follows: work at sites 4, 1, 3, and 2 in Texas, followed by work at the sites in Oklahoma, Arkansas, and Colorado.

The characteristics of the sites are summarized in Table 3. All sites had 12-ft lanes with paved shoulders. The only visible markings were centerlines made of yellow reflective tape. Annual average daily traffic rates (AADTs) ranged between 2,530 and 6,700 vehicles. Three of the sites were located in highway sections with relatively level terrain, and four sites were in sections with rolling terrain. All of the sites included a horizontal curve and a tangent section. The degree of curvature was 2.0 degrees at six sites and 3.0 degrees at one site. Some of the sites included sections that would be marked again as no-passing zones after the pavement overlay construction work was completed.

Operational Measurements

Traffic stream measurements included vehicle speed, lateral distance from the centerline (measured from the centerline to the outer edge of the left front tire), lane encroachment (straddling centerline), and erratic maneuvers (e.g., abrupt swerving, excessive slowing, stopping, etc.). Vehicle speed, lateral distance, and lane encroachment data were collected by using an automated data collection system developed by TTI. Tapeswitches attached to the pavement surface were wired to computers housed in vehicles that were parked off the roadway as far from the operating lanes as possible.

A schematic of the tapeswitch placements and data collection configuration for a typical field study is shown in Figure 1. The specifics concerning the installation were as follows:

- One Z-type tapeswitch configuration was installed at a base station located upstream of the test section to record times, speeds, lateral distances, and encroachments of vehicles on a roadway section containing the highway agency's existing temporary centerline pavement marking pattern.
- Three Z-type tapeswitch configurations were located in the curve section at the 1/4, 1/2, and 3/4 distance points from the beginning of the curve.
- Three Z-type tapeswitch configurations, spaced about 400 ft apart, were located in a tangent section.
- One double tapeswitch configuration was located in the opposing lane near the curve tapeswitches and one near the tangent tapeswitches. The double tapeswitches recorded the times and speeds of opposing vehicles.

In addition to data recorded with the automated system, field personnel located near the two computer systems observed erratic maneuvers within the horizontal curve and tangent sections.

In-Vehicle Driver Response

In-vehicle studies were conducted to supplement the traffic stream evaluation. Each of 27 paid driver subjects (four at each of six sites and three at one site), recruited from the local areas, was accompanied by a TTI study administrator as he or she drove through one of the seven test sites. Each subject drove through a study site on each of three nights while traffic stream data were being collected for the 1-ft, 2-ft, and 4-ft striping patterns. The administrator recorded driver comments and erratic maneuvers and administered a post-drive-through survey each night to obtain additional information. Details of the survey forms and instructions can be found elsewhere (2). Age and gender distributions of the driver subjects are presented in Table 4 in relation to a national distribution of drivers (8). The first number shown is the proportion needed to match the national demographic in age and gender.

A speed/distance recorder, used at the four Texas sites, provided information necessary for developing driver speed profiles. Electronic problems in the test vehicle prevented TTI researchers from recording similar data in Oklahoma, Arkansas, and Colorado.

Experimental Design

Each marking pattern was tested at each site on consecutive week nights except when inclement weather or equipment

TABLE 3 STUDY SITE CHARACTERISTICS

Site	State	Route	Location	AADT	Study Direction	Section Length (ft)	Curve Length (ft)	Curve Direction	Degree of Curve	Lane/ Shoulder Width (ft)	Terrain
1	Texas	US 190	1 mi north of Milano	5,200	Northbound	6,700	1,022	RH	2.0	12/10	Level
2	Texas	SH 36	1 mi west of Brenham	9,600	Southbound	3,700	1,600	LH	2.0	12/10	Rolling
3	Texas	SH 276	1.5 mi east of Rockwall	6,000	Eastbound	2,530	1,831	RH	2.0	12/10	Level
4	Texas	US 96	Silsbee Bypass	5,000	Southbound	3,880	1,850	RH	2.0	12/10	Level
5	Oklahoma	US 64	Eastern edge of Sallisaw	3,000	Eastbound	3,640	1,060	RH	2.0	12/5	Rolling
6	Arkansas	US 71	4.5 mi north of Wickes	3,750	Northbound	3,200	700	LH	3.0	12/4	Rolling
7	Colorado	US 160	27 mi east of Durango	2,750	Eastbound	3,000	1,260	RH	2.0	12/5	Rolling

NOTE: All sites were at overlay projects on two-lane, two-way rural highways. The centerline stripes were the only markings on the highway sections.

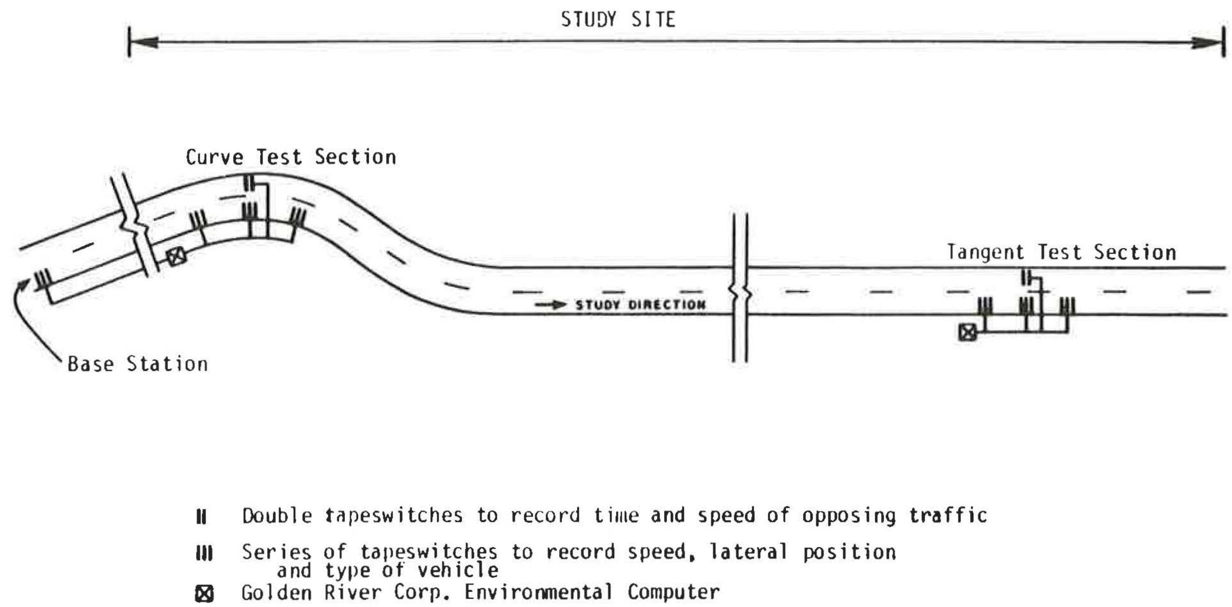

FIGURE 1 Data collection configuration.

TABLE 4 AGE AND GENDER DISTRIBUTION OF SUBJECT DRIVERS

	Age[a]					
	≥ 24	25–34	35–44	45–54	55 ≥	Total
Males	3/4	3/3	3/2	2/2	3/3	14/14
Females	2/3	3/1	3/3	2/3	3/3	13/13
Total	5/7	6/4	6/5	4/5	6/6	27/27

[a]Number of subjects needed/number of subjects tested.

problems prevented testing. The pavement tape used to mark the test centerline was manually removed each day and replaced with a new striping pattern in time for the nighttime studies. Removal of the stripes each day did not leave any visible markings on the pavement, regardless of the order in which the patterns were studied.

The order of striping patterns is shown in Table 5. The order of treatments was counterbalanced across sites according to a Latin square design. Note that each treatment was scheduled at least twice in each order position. Because there were seven sites rather than six, one treatment appeared three times in each order position. In regard to field data, two exceptions to the original counterbalanced design were required. Weather problems required an adjustment in the order during the studies at Site 7 (Colorado). The high potential of having to abandon the

TABLE 5 ORDER OF STRIPING PATTERN TESTS

	Pattern, ft		
Site	First Night	Second Night	Third Night
1	2	1	4
2	4	2	1
3	1	4	2
4	4	2[a]	1
5	1	4	2
6	2	1	4
7	4	1	2

[a]No data available.

studies at Site 7 because of prolonged inclement weather led to a decision to study the 1-ft stripe (instead of the 2-ft stripe) immediately after studying the 4-ft stripe. Loss of data for the 1-ft stripe was considered to be more critical than the loss of data for the 2-ft stripe. The rationale was that comparisons could be made between the two extreme test striping patterns. If no differences were found between the 1-ft and 4-ft patterns (as had been the case in the Texas studies), then it could be concluded that there would be no differences between the 2-ft and 4-ft patterns. The weather, however, did clear long enough to collect data for the 2-ft stripe after data were collected for the 1-ft stripe.

At Site 4, the initial study site, field data on the second night were lost due to inclement weather and equipment problems. The subject questionnaire was administered under all treatments, and the lower ratings at Site 4 on the second night may be expected to partially reflect the inclement weather.

Analysis Approach

Practical Speed and Lateral Distance Differences

Because of the large sample size expected from the field studies, it was anticipated that statistical significance would be detected in even small differences in average speeds and lateral distances between the 1-ft, 2-ft, and 4-ft treatments. The concern of the research team was to ensure that the results would be interpreted not only statistically but also from a practical standpoint. For example, during analysis of the differences between two of the temporary pavement marking patterns, a difference in average speeds of 1 mph might be found to be statistically significant because of the large sample size. However, from a practical standpoint, a 1-mph speed difference would be rather meaningless. It therefore became necessary to identify a speed differential that would be considered acceptable in a practical sense. On the basis of the many years of research and operational experience of the research team and discussions with several other traffic safety and operations experts, a speed difference of 4 mph or greater was chosen as practically significant. Similarly, the team chose differences in lateral distance of 1 ft or greater occurring at four of the six curve and tangent sensor stations as practically significant.

Tracing Vehicles: A More Powerful Analysis Design

Another important feature of the analysis was the analysis experimental design. The ability to trace individual vehicles from the base sensor station through each of the other six sensor stations (three in the curve and three in the tangent) allowed the use of a matched or paired comparison statistical design that significantly increased the power of detecting significant differences among the 1-ft, 2-ft, and 4-ft striping patterns. This increased power translated into a reduced sample size requirement for detection of differences with the same precision as an unmatched design (vehicles are not traced through the sensor stations). For example, a sample size of 63 matched (traced) vehicles would be equivalent to 125 unmatched vehicles (about 2 times as many) when detecting average speed differences of 4 mph at the 0.05 level of significance and 80 percent power.

FIELD STUDY RESULTS

This section of the paper discusses the combined results from the seven field study sites. Details for each study site are presented elsewhere (2).

Sample Size

Vehicle sample sizes by site and pavement marking pattern are presented in Table 6. The sample sizes are listed in four groups: (a) the total number of vehicles observed during the studies, (b) the number of traced vehicles, (c) the total number of vehicles with headways of 4 sec or longer, and (d) the number of traced vehicles with headways of 4 sec or longer. A total of 3,697 vehicles were sampled at the seven overlay study sites. Of these vehicles, 2,883 were traced through at least the base and the three curve stations (2,814 were traced through all seven stations), 2,803 had headways of 4 sec or longer, and 2,518 vehicles with headways of 4 sec or longer were traced through at least the base and the three curve stations (2,443 vehicles with headways of 4 sec or longer were traced through all seven stations).

Traced Vehicles with Headways > 4 Sec

For each study site, Tables 7 and 8 summarized the average speeds and the average lateral distances from the centerline at the base and the three curve and three tangent sensor stations. The curve stations, CURVE-1, CURVE-2, and CURVE-3, were located at the 1/4, 1/2, and 3/4 distance points from the beginning of the horizontal curve. The tangent stations, TAN-1, TAN-2, and TAN-3, were located at 400-ft spacings, with the exception of Site 7 (Colorado). The spacing at Site 7 was reduced to 250 ft because the available tangent section was short.

TABLE 6 SAMPLE SIZE

Stripe	Site 1 (TX)	Site 2 (TX)	Site 3 (TX)	Site 4 (TX)	Site 5 (OK)	Site 6 (AR)	Site 7 (CO)	TOTAL
TOTAL NUMBER OF VEHICLES OBSERVED								
1-ft	125	184	349	43	118	146	138	1,103
2-ft	170	123	659	**	169	150	100	1,371
4-ft	148	192	434	38	125	149	137	1,223

								3,697
NUMBER OF TRACED VEHICLES								
1-ft	110	79	313	43	82	125	82	834
2-ft	137	97	611	**	116	95	73	1,129
4-ft	89	72	415	38	89	111	106*	920

								2,883
NUMBER OF VEHICLES WITH HEADWAYS ≥ 4 SECONDS								
1-ft	112	150	294	43	105	126	91	921
2-ft	143	99	524	**	129	95	76	1,066
4-ft	89	68	361	38	109	104	47	816

								2,803
NUMBER OF TRACED VEHICLES WITH HEADWAYS ≥ 4 SECONDS								
1-ft	106	62	271	43	80	114	73	749
2-ft	130	83	488	**	104	84	70	959
4-ft	84	58	344	38	84	97	(105)*	810

								2,518

*There were a maximum of 106 (105) vehicles traced through the curve only; 37 (30) vehicles were traced through all 7 stations.

**No data available.

TABLE 7 AVERAGE SPEEDS IN MPH FOR TRACED VEHICLES WITH HEADWAYS ≥ 4 SEC

N	STRIPE	BASE	CURVE-1	CURVE-2	CURVE-3	TAN-1	TAN-2	TAN-3
SITE 1 - TX								
106	1-ft	61	58	58	60	59	59	**
130	2-ft	60	58	59	61	59	60	**
84	4-ft	62	58	58	60	59	59	**
SITE 2 - TX***								
62	1-ft	56	56	56	**	56	57	56
83	2-ft	56	56	57	55	56	55	53
58	4-ft	57	57	58	57	57	57	55
SITE 3 - TX								
271	1-ft	57	54	53	53	**	**	**
488	2-ft	57	54	54	53	**	**	**
344	4-ft	58	55	56	54	**	**	**
SITE 4 - TX								
43	1-ft	**	56	55	56	**	**	**
0	2-ft	**	**	**	**	**	**	**
38	4-ft	**	59	57	58	**	**	**
SITE 5 - OK								
80	1-ft	58	55	55	53	55	56	56
104	2-ft	56	54	54	53	55	55	55
84	4-ft	56	53	52	52	53	54	54
SITE 6 - AR								
114	1-ft	57	58	56	54	55	56	54
84	2-ft	56	56	55	53	54	56	54
97	4-ft	57	57	56	55	54	55	54
SITE 7 - CO***								
73	1-ft	53	51	53	53	51	51	51
70	2-ft	54	**	53	54	52	52	52
105	4-ft	54	53	54	55	54	54	55

**No data available.

***Tangent data collection stations preceded the curve stations.

The statistical analysis revealed that there were no significant differences in average speed among the 1-ft, 2-ft, and 4-ft striping patterns, with the exception that statistically significant differences were found at the Site 1 CURVE-2 sensor location and at the Site 3 CURVE-2 and CURVE-3 sensor locations. However, the average speed differences at these three sensor stations were 2.5 mph or less and were not considered to be practically significant.

The analysis of variance procedure assumes that the variability among treatment groups is homogeneous. This assumption was tested using the Sheffé F-test for comparing population variances. Basically, as a rule of thumb, if sample standard deviations are within a factor of 2 (i.e., if the minimum sample standard deviation doubled does not exceed the maximum sample standard deviation), then the variance for all groups can be considered statistically equal. A review of the data revealed that the standard deviations at only two stations at one site (Site 2) were greater by a factor of 2. Therefore it was felt that the assumption of homogeneous variance is valid for these data.

Analysis of the lateral distance data also revealed that there were no statistical or practical differences in lateral distance from the centerline among the three striping patterns. The average lateral distance differences between the 1-ft and 2-ft striping patterns and the 4-ft striping pattern were only 0.4 ft or less.

Analysis of the vehicle encroachment (straddling) data did not reveal any patterns either. Centerline encroachment at the sensor stations was extremely infrequent and sporadic. The field observers noted that the few cases of vehicle encroachment that did occur were due to passing maneuvers and other factors unrelated to the centerline striping pattern. Erratic maneuvers caused by the striping patterns were also essentially nonexistent.

As additions to the data just mentioned, speed profiles of the subject drivers at the four Texas sites were developed to determine whether the speed patterns could help distinguish differences between the three pavement marking patterns. The speed profile sample size at each site (3 or 4 subjects) was not

TABLE 8 AVERAGE LATERAL DISTANCES IN FT FROM CENTERLINE FOR TRACED VEHICLES WITH HEADWAYS ≥ 4 SEC

N	STRIPE	BASE	CURVE-1	CURVE-2	CURVE-3	TAN-1	TAN-2	TAN-3
SITE 1 - TX								
106	1-ft	4.0	4.6	4.2	4.9	3.3	3.6	**
130	2-ft	4.0	4.6	4.4	5.2	3.5	3.6	**
84	4-ft	4.3	4.9	4.7	5.3	3.7	4.0	**
SITE 2 - TX***								
62	1-ft	4.8	3.2	3.1	**	3.3	2.7	2.6
83	2-ft	4.2	3.2	3.0	2.9	3.7	2.8	2.8
58	4-ft	3.6	3.1	2.9	2.8	3.0	2.9	2.6
SITE 3 - TX								
271	1-ft	4.4	4.7	5.0	4.8	**	**	**
488	2-ft	3.8	4.8	4.7	3.7	**	**	**
344	4-ft	4.1	4.9	5.2	4.0	**	**	**
SITE 4 - TX								
43	1-ft	**	5.5	4.7	6.3	**	**	**
0	2-ft	**	**	**	**	**	**	**
38	4-ft	**	5.8	4.9	6.1	**	**	**
SITE 5 - OK								
80	1-ft	2.6	3.5	3.3	4.8	3.6	3.2	4.4
104	2-ft	2.7	3.6	3.4	5.1	3.3	3.0	4.1
84	4-ft	2.6	3.6	3.5	5.1	3.6	3.0	3.7
SITE 6 - AR								
114	1-ft	2.1	2.3	2.3	2.0	1.9	2.2	2.0
84	2-ft	2.0	2.4	2.0	1.9	2.0	2.0	1.7
97	4-ft	1.8	2.2	2.1	1.9	2.1	2.5	1.9
SITE 7 - CO***								
73	1-ft	4.1	3.0	4.0	3.2	2.5	2.4	3.2
70	2-ft	2.7	3.2	**	3.0	2.3	2.2	2.7
105	4-ft	3.0	3.9	2.8	3.4	2.4	2.0	2.6

*Measured from the centerline to the outer edge of the left front tire.

**No data available.

***Tangent data collection stations preceded the curve stations.

large enough for the performance of statistical analyses. However, visual inspection of the speed profiles revealed no consistent speed patterns that indicated any differences between the three striping patterns. The speed profiles showed considerable variability and seemed to be indicators of individual driving habits rather than the results of differences among the three striping patterns.

In summary, speed, lateral distance, encroachment, erratic maneuver, and speed profile data for the sample of vehicles with headways of 4 sec or more indicated that there were no differences in driver performance between the 1-ft, 2-ft, and 4-ft striping patterns.

Subject Evaluations

Table 9 presents the ratings by the subject drivers of the three pavement marking treatments across the seven study sites. Each driver was asked to rate the markings on a scale as follows: 4- extremely effective, 3- effective, 2- satisfactory, 1- not very effective, and 0- unsuitable and possibly dangerous. With one exception, there were four driver/subjects per site, so the maximum total rating per striping pattern per site was 16. If all four subjects judged the treatment pattern effective, the total rating was 12. If all rated it satisfactory, the total rating was 8.

Site 1 had only three subjects. To include these data (shown in Table 9 within parentheses) with the other data, it was necessary to extrapolate the ratings as if there had been four subjects. The same procedure was followed with the ranking data.

The results showed that few drivers used the 4 rating and no one used the 0 rating. The average rating across all studies was a 10.8, slightly below the 12.0 (effective) rating. At two sites (2 and 7), mean ratings were 12.7 across treatments, whereas three sites had mean ratings of 9.3 to 9.7. There appeared to be a slight trend toward a relationship between order of effectiveness and length of the stripe. At only one site was the 1-ft stripe

TABLE 9 SUMMARY OF RATINGS OF PATTERNS AT SEVEN SITES

Treatment	Site 1[a] (Tex.)	Site 2 (Tex.)	Site 3 (Tex.)	Site 4 (Tex.)	Site 5 (Okla.)	Site 6 (Ark.)	Site 7 (Colo.)	Total	Mean
1 ft	9.3 (7)[b]	12	7	12[c]	7	8	13	68.3	9.8
2 ft	12 (9)[b, c]	13[c]	11[c]	8	11[c]	9	11	75.0	10.7
4 ft	10.4 (8)[b]	13[c]	11[c]	10	10	12[c]	14[c]	80.4	11.5
Mean	10.8	12.7	9.7	10.0	9.3	9.7	12.7	223.7	10.8

CODE: 16 = extremely effective, 12 = effective, 8 = satisfactory, 4 = not very satisfactory. Max = 16, min = 0.
[a]Site 1 had only three subjects. Extrapolations were made on the basis of four subjects for comparisons of ratings.
[b]Original rating based on three subjects.
[c]Best rating.

judged to be most effective. At two sites there was a strong preference for the 4-ft stripe, but overall, subjects lacked a strong preference between the 2- and 4-ft striping patterns. The variability across studies led to no significant difference between ratings.

The data in Table 10 summarize the ranking data across studies. A ranking of most effective was assigned a 1, next most effective a 2, and least effective a 3. Hence the best possible ranking at a site was a 4, and the poorest possible rank a 12. The mean ranking for each marking pattern across studies varied only slightly from the mean of 8.0. Again, the 1-ft stripe was poorest (9.2), but it was not significantly different from the 2-ft and 4-ft stripes.

After the studies at Sites 1 and 4 (Texas), the subject questionnaire was modified for the next five sites in an attempt to assess whether the drivers were even aware that there were differences in the three striping patterns. Drivers were instructed after the third and final night to rank the stripes in length, in spacing between them, and in brightness. They were not told in advance to look for these particular features, but it was important to know whether the drivers were basing the effectiveness ratings and rankings on some design feature rather than on extraneous factors unique to the site, such as the weather or the traffic.

Table 11 shows that drivers at all five sites ranked the 4-ft stripe as being longer than the 1-ft stripe. However, the drivers at Sites 2 and 5 had difficulty in discriminating the 2-ft and 4-ft stripes, and drivers at Sites 6 and 7 could not distinguish differences in the 1-ft and 2-ft lengths. In general, there was a trend toward being able to distinguish differences in length, even though the drivers were not asked to do so in advance.

Table 12 presents the estimates of spacing between stripes. Drivers at all but one of the sites reported that the 39-ft spacing (1-ft stripe) was greater than either the 38-ft spacing (2-ft stripe) or 36-ft spacing (4-ft striping). Strangely, they could not discriminate between the 38- and 36-ft spacing even though there was a 2-ft difference, while the 1-ft difference was detected.

Table 13 shows the estimates of brightness. Subjects at two sites were convinced that the 2-ft stripe was brightest and the 1-ft was the dimmest, but at the other three sites there was virtually no difference. Overall, drivers could not discriminate among brightness levels.

Drivers' Comments

Drivers' comments were highly variable and often dwelled on situational factors unrelated to the pavement marking patterns.

TABLE 10 SUMMARY OF RANKINGS OF PATTERNS AT SEVEN SITES

Treatment	Site 1[a] (Tex.)	Site 2 (Tex.)	Site 3 (Tex.)	Site 4 (Tex.)	Site 5 (Okla.)	Site 6 (Ark.)	Site 7 (Colo.)	Total	Mean
1 ft	10.7 (8)[b]	7[c]	12	6.5[c]	11	7[c]	10	64.2	9.2
2 ft	7.3 (5.5)[b]	7[c]	6.5	7	6[c]	9	7[c]	49.8	7.1
4 ft	6 (4.5)[b, c]	10	5.5[c]	7.5	7	8	7[c]	51.0	7.3
Mean	8.0	8.0	8.0	7.0	8.0	8.0	8.0	165.0	7.9

CODE: 4 = best, 8 = second best, 12 = worst. Max = 4, min = 12.
[a]Site 1 had only three subjects. Extrapolations were made on the basis of four subjects for comparisons of ratings.
[b]Original rating based on three subjects.
[c]Best rating.

TABLE 11 STRIPE LENGTH ESTIMATES AT FIVE SITES

Treatment	Site 2 (Tex.)	Site 3 (Tex.)	Site 5 (Okla.)	Site 6 (Ark.)	Site 7 (Colo.)	Total	Mean
1 ft	5	5	5	7	7	29	5.8
2 ft	10[a]	7	10[a, b]	7	7	41	8.2
4 ft	9	12[a]	10[a, b]	11[a]	11[a]	53	10.6

CODE: 3 = longest (12 max), 2 = midlength (8 mid), 1 = shortest (4 min). Note that questions of length, spacing, and brightness were not asked in first two studies. Mean rank across treatments is 8 by procedure.
[a]Longest stripe.
[b]One tie for longest.

TABLE 12 SPACING LENGTH ESTIMATES AT FIVE SITES

Treatment	Site 2 (Tex.)	Site 3 (Tex.)	Site 5 (Okla.)	Site 6 (Ark.)	Site 7 (Colo.)	Total	Mean
1 ft (39-ft space)	10[a]	12[a]	12[a]	8	9[a]	51	10.2
2 ft (38-ft space)	8	5	5[b]	8	8	34	6.8
4 ft (36-ft space)	6	7	6[b]	8	7	34	6.8

CODE: 3 = longest space (12 max), 2 = midlength space (8 mid), 1 = shortest space (4 min).
[a]Longest space.
[b]One tie for least spacing.

TABLE 13 STRIPE BRIGHTNESS ESTIMATES AT FIVE SITES

Treatment	Site 2 (Tex.)	Site 3 (Tex.)	Site 5 (Okla.)	Site 6 (Ark.)	Site 7 (Colo.)	Total	Mean
1 ft	9[a]	5	6	9[a]	8	37	7.4
2 ft	8	11[a]	10[a]	7	8	44	8.8
4 ft	7	8	8	8	8	43	8.6

CODE: 3 = brightest (12 max), 2 = midbrightness (8 mid), 1 = dimmest (4 min).
[a]Brightest.

However, when comments on length, spacing, brightness, or effectiveness were volunteered, these comments were generally reflected in the drivers' rankings, ratings, and estimates.

To summarize, the 1-ft, 2-ft, and 4-ft patterns were all rated satisfactory to effective. There was no statistical difference in ratings, but the trend was toward judging the 1-ft stripe as less effective (only at one site was the 1-ft stripe judged most effective). There was no difference in rankings, but again, the trend was toward the 1-ft stripe being ranked as slightly poorer. At four of seven sites, it was ranked as much poorer than the other two lengths.

Drivers were able to distinguish the lengths of the 4-ft and 1-ft stripes but had difficulty distinguishing between the 1- and 2-ft lengths and between the 2- and 4-ft lengths. They could tell that the 39-ft spacing, associated with the 1-ft stripe, was the longest, but they could not tell the difference between the 38-ft and 36-ft spacings. Brightness estimates were virtually random. Had the drivers been instructed in advance to concentrate on these features or if the patterns had been viewed successively on the same night, performance might have been better. However, to have done so might have biased the drivers toward basing their effectiveness and ranking judgments on these features.

After studies at three field sites in Texas, the subject questionnaire was modified to obtain direct statements from the subjects about the adequacy of the 1-ft, 2-ft, and 4-ft striping patterns. The following question was added and asked each night after the drive through: "Does this marking pattern provide adequate path delineation?"

The responses to this question are summarized in Table 14. From the table, it can be seen that 13 of the 16 subjects interviewed stated that the 1-ft striping pattern did provide adequate delineation, 15 stated that the 2-ft striping pattern was adequate, and all 16 believed that the 4-ft striping pattern was adequate. In general, drivers slightly preferred the longer stripes, but there is no compelling evidence that the 2- or 4-ft stripes are superior to the 1-ft stripe.

TABLE 14 SUBJECTS STATING THAT STRIPING PATTERN PROVIDED ADEQUATE DELINEATION–FOUR SITES

Treatment	Site 1 (Tex.)	Site 5 (Okla.)	Site 6 (Ark.)	Site 7 (Colo.)	Total
1 ft	3	3	3	4	13
2 ft	4	4	3	4	15
4 ft	4	4	4	4	16

FINDINGS AND SUGGESTED RESEARCH

On the basis of driver performance and driver subjective evaluations, the 1-ft and 2-ft on 40-ft centers striping patterns performed as well as the 4-ft pattern for centerline striping at night at seven pavement overlay projects on rural two-lane, two-way highways with 2.0 degree horizontal curvatures, level to rolling terrain, and average speeds between 50 and 62 mph. Studies were conducted in four states.

The findings should not be generalized to situations not tested. Nighttime viewing in an ambient background of near darkness will enhance the contrast of the bright reflective yellow stripes. Moreover, the horizontal curves were 2.0 degrees, with the exception of one curve that was 3.0 degrees. It is possible that the performance of the three tested striping patterns may not be equal on horizontal curves with greater curvature or at urban or suburban construction zones where the ambient lighting is different than the conditions studied. Also, the three striping patterns tested may not result in the same driver performance on mountainous highways and other types of highways with lower operating speeds.

The study did not attempt to optimize spacing or brightness to determine the most cost-effective striping pattern. Although the three striping patterns tested provided adequate delineation on rural two-lane, two-way highways, they may not necessarily represent the optimum patterns from a cost-effectiveness standpoint. It is possible that patterns with larger spacings may also provide adequate path delineation on rural two-lane, two-way highways.

The limitations of this research relative to scope of the field studies were discussed in the previous section. The discussion suggests the following:

- Future research should be directed at the effectiveness of the three striping patterns (1, 2, and 4 ft on 40-ft centers) at construction zones in situations with less brightness contrast (suburban and urban areas), horizontal curvatures greater than 2 degrees, mountainous terrain, and operating speeds lower than those tested in the current study. Ideally, studies should also be conducted when the pavement is wet and during rain.
- Research should also be directed at determining the optimum spacing of the 1-ft, 2-ft, and 4-ft stripes at construction zones on two-lane, two-way rural highways.

ACKNOWLEDGMENTS

The study reported in this paper was part of NCHRP Project 3-32, "Temporary Pavement Markings for Work Zones." Appreciation is expressed to the following individuals for their guidance during the project and review of the final report for the project: Robert Spicher and Dan Rosen, TRB; Harold R. Hofener, Oklahoma DOT and chairman of the Project 3-32 Panel; Gerson Alexander, Positive Guidance Applications, Rockville, Maryland; Richard F. Beaubien, City of Troy, Michigan; Linn D. Ferguson, California DOT (Caltrans); Woodrow H. Gump, Illinois DOT; Robert S. Hostetter, Institute for Research, State College, Pennsylvania; Robert A. Kurpius, Minnesota DOT; Arthur W. Roberts III, New Jersey DOT; and Justin True, FHWA.

REFERENCES

1. *Federal Register*, Vol. 47, No. 24, Feb. 4, 1982, p. 5245.
2. C. L. Dudek, R. D. Huchingson, and F. T. Creasey. *Temporary Pavement Markings for Work Zones.* NCHRP Project 3-32. Feb. 1987.
3. H. Godthelp and J. B. J. Riemersma. *Perception of Delineation Devices in Road Work Zones During Nighttime.* Preprint. Society of Automotive Engineers, Warrendale, Pa., 1982.
4. H. Godthelp and J. B. J. Riemersma. Vehicle Guidance in Road-Work Zones. *Ergonomics*, Vol. 25, No. 10, Oct. 1982, pp. 909–916.
5. L. J. Spencer. *Construction Zone Delineation: Raised Pavement Markers.* Final Report. Arkansas State Highway and Transportation Department, Little Rock, March 1978.
6. C. W. Niessner. *Construction Zone Delineation (Raised Pavement Markers).* Report FHWA-TS-78-222. FHWA, U.S. Department of Transportation, 1978.
7. C. L. Dudek, R. D. Huchingson, and D. L. Woods. Evaluation of Temporary Marking Patterns for Work Zones. In *Transportation Research Record 1086*, TRB, National Research Council, Washington, D.C., 1980, pp. 12 –20.
8. *Highway Statistics 1984*, Federal Highway Administration, U.S. Government Printing Office, Washington, D.C.

DISCUSSION

ANITA W. WARD

Highway Products Division, Potters Industries Inc., 20 Waterview Boulevard, Parsippany, N.J. 07054

Within the scope and test conditions specified by NCHRP, the Texas Transportation Institute (TTI) chose a set of near-perfect conditions. On the pavement overlay projects on rural two-lane, two-way highways, the following conditions prevailed:

- Site locations were relatively level or rolling terrain.
- Although all sites included a horizontal curve and a tangent section, the degree of curvature at six sites was only 2.0 degrees; at the seventh site, it was 3.0 degrees.
- All sites had 12-ft lanes with 4-ft to 10-ft paved shoulders.
- Site speeds averaged between 50 and 62 miles per hour.
- The ambient background was "near darkness."
- The marking material selected for this field test was highly retroreflective pavement-marking tape.

These last two points have a major impact on skewing the research results. Drivers actually see pavement markings as a function of their contrast with the road surface (*1*). New pavement overlays such as those in the test sites are generally very black, providing excellent contrast to the yellow marking tape. The retroreflective properties of the newly applied marking tape itself provide an extremely bright optical target. On a clear, dry night (each one in this study), it is far easier for a driver to see such highly visible pavement markings than in most driving situations. The only visual distractions appear to have been limited traffic and the data collection system of "computers housed in vehicles parked off the roadway as far from the operating lanes as possible." Note that there was no discussion of potential change in driver behavior as a result of the parked vehicles.

One condition imposed by NCHRP is likely to have further skewed TTI's results: the absence of edgelines. Experience and an ample body of evidence indicate improved driver performance in the presence of edgelines (*2*). With no indication of lane boundary and limited visual information at the edge of pavement in this study, drivers' focus on the centerline was even more acute.

With a wide expanse of blacktop in a nearly dark environment and a brilliant ribbon of yellow to follow, as in this study, drivers should perform relatively consistently. It is not surprising that TTI's summary of traced vehicles indicates no differences in driver performance between the 1-ft, 2-ft, and 4-ft striping patterns with the measurement criteria of speed, lateral distance encroachment, erratic maneuver, and speed profile data. Similarly, it is reasonable to expect that given the strong visual target of highly retroreflective new pavement-marking tape contrasted against a newly surfaced road in a background of near darkness, some individuals in the subject evaluations could not differentiate between the 1-ft and 2-ft stripes, or even perhaps between 2-ft and 4-ft stripes. Each is perceived as a very bright spot in a black environment. Such spots may also appear elongated by the relatively high speeds. The TTI observation that drivers could differentiate the 39-ft spacing separating the 1-ft spots of bright light but could not discriminate between the 38-ft and 36-ft spacing separating the 2- and 4-ft bright spots supports this.

Perhaps the most surprising result of the field study was that even with these ideal conditions, *each method* of subject evaluation reported the poorest results with the 1-ft stripe and a preference for the 4-ft stripe. Yet far more important is the reported finding that *none* of the treatments were judged as extremely effective. The treatments were only rated 2 on a scale of 0 to 4. This is consistent with a prior TTI research study that reported that drivers rated 8-ft stripes with 32-ft gaps as the best striping treatment (*3*).

Given all this, it is imperative that the data in this field study are not interpreted as representative of a pavement marking pattern. They can at best be indicative of a newly placed, highly retroreflective pavement-marking tape on a resurfaced road where there is little or no visual "clutter."

The "typical" construction zone does not meet test conditions selected for TTI's study, and work zone safety is becoming a more critical issue. Analysis of U.S. traffic accidents reveals that work zone fatalities have increased from 490 in 1982 to 680 in 1985 (*4*). The Standing Committee on Highway Traffic Safety of the American Association of State Highway and Transportation Officials conducted a survey of work zone accidents on the Interstate and Primary System in 1985. Their summary reported (*4*) that

- There was an estimated $800 million economic cost associated with 400 fatal accidents, 15,000 injury accidents, and 31,000 property damage–only accidents.
- Work zone fatal accidents are concentrated in rural areas.
- Work zone accidents produce more injuries and fatalities than the national average for all accidents.
- Although more than two-thirds of all accidents occur in daylight, nighttime accidents are far more severe. Nighttime accidents account for more than half of the fatal accidents and more than their share of injury accidents.

Work zones are particularly hazardous because they present drivers with changes in the normal driving environment. Such changes place greater demand on drivers, possibly leading to confusion and accidents. Up to 90 percent of all the information used by drivers to guide and control their vehicles is obtained visually (*5*), and the pavement itself is a primary information source for drivers. In fact, if drivers are presented with conflicting information, they will generally choose to follow the pavement (*6*).

Pavement markings through work zones should provide a clear path for drivers' guidance. Such markings must be effective where needed most: at night, under adverse weather conditions, and when drivers may have other visual limitations, such as advancing age, fatigue, or alcohol consumption. The need for strong delineation patterns in work zones is gaining widespread acceptance, and our court system is providing impetus for action. In both Louisiana and New Mexico, the states were held liable for wrongful deaths where striping was not in place to warn and guide motorists through work zones (*7*). The state of North Carolina has taken the lead in providing increased information through construction work zones by using 8-in. markings, twice the standard marking line width (from a letter by J. M. Lynch to W. Cromartie, North Carolina DOT, Raleigh, August 8, 1985).

Safe driving requires both appropriate visual information and drivers who are able to receive and interpret that information. However, studies indicate that in most construction zone accidents, the driver receives neither visual stimulation nor sufficient warning (*8*). The fact that drivers often fail to meet the challenges of work zones is documented by studies indicating that the accident rate increases in work zones during construction, as compared to before construction (*8–10*). Drivers cannot effectively control their vehicles without sufficient visual information, and even this current TTI study indicates that a pavement marking pattern of short stripes with long gap ratios does not provide an effective level of communication. There is a significant body of evidence to indicate that driver performance is enhanced through stronger pavement marking patterns (*11*).

The negative consequences of this report could be far-reaching. Even though the report states "that the findings should not be generalized to situations not tested," response to this presentation at the annual Transportation Research Board meeting indicates that this is precisely what will happen. The potential detrimental impact to safety and mobility is heightened by TTI's own conclusions: With evidence only of *treatment* (with highly retroreflective marking tape under ideal conditions), TTI has claimed not only that the striping *patterns* of 1 ft, 2 ft, and 4 ft on 40-ft centers are adequate but that even larger spacings may help to optimize cost effectiveness.

As indicated in the statement of the problem, TTI uses a very narrow interpretation of the word "cost." Cost is not just money spent. More importantly, cost is measured in value received. If drivers cannot safely position their vehicles through a work zone to prevent harm to those individuals or objects in the area and to protect themselves and their passengers, a responsible jurisdiction should not open that area to traffic. Sound business considerations and concern for the public welfare dictate comprehensive resource management. Inadequate pavement marking patterns, especially in work zones where drivers need enhanced visual communication, are a prime example of false economy.

REFERENCES

1. *Visual Aspects of Road Markings*. Final Draft. CIE Division 4, Technical Committee 4.04, Permanent International Association of Road Congresses, 1986.
2. M. Anton-Thomas. Surviving the Test of Time. Presented at the Institution of Transportation Engineers Annual Meeting, Indianapolis, Ind., 1986.
3. C. L. Dudek, et al. Evaluation of Temporary Pavement Marking Patterns in Work Zones: Proving-Ground Studies. In *Transportation Research Record 1086*, TRB, National Research Council, Washington, D.C., 1986, pp. 12–20.
4. *Summary Report on Work Zone Accidents*. Standing Committee on Highway Traffic Safety, AASHTO, Washington, D.C., 1987.
5. M. J. Allen. *Vision and Highway Safety*, Chilton, Philadelphia, 1970.
6. R. M. Lewis. Safer Work Zone Devices: The Next Generation. *American Transportation Builder*, Vol. 62, No. 4, 1985, p. 10.
7. R. W. Anderson. On the Road. *Transafety Reporter*, Vol. 5, No. 5, 1987, pp. 4–5.
8. R. F. Pain et al. *NCHRP Report 236: Evaluation of Traffic Controls for Highway Work Zones*. TRB, National Research Council, Washington, D.C., 1981.
9. J. L. Graham et al. *Accident and Speed Studies in Construction Zones*. Report FHWA-RD-77-80. FHWA, U.S. Department of Transportation, 1977.
10. Z. A. Nemeth and D. J. Migletz. *Accident Characteristics Before, During, and After Safety Upgrading Projects on Ohio's Rural Interstate System*. Ohio Department of Transportation, Columbus, 1977.
11. J. A. Deacon. Potential Improvements in Roadway Delineation for Older Drivers. In *Special Report 218*. TRB, National Research Council, Washington, D.C., 1988.

AUTHORS' CLOSURE

We do not agree with the discussant's claim that the design of this research project led to questionable conclusions. Our response will show that the methodology and analysis were in fact sound and led to valid conclusions.

The experimental question was simply, Does the length of the temporary stripe (4, 2, or 1 ft) and associated 36-, 38-, and 39-ft gaps make a difference in how motorists drive when temporary pavement markings are used on overlay projects on two-lane, two-way rural highways? Field studies were conducted at seven pavement overlay project sites on two-lane, two-way rural highways in Arkansas, Colorado, Oklahoma, and Texas to determine whether motorists would drive better with 4-ft stripes. The data failed to indicate major differences. The in-car studies with paid drivers were included to provide a medium in which drivers could express their opinions; these were taped in real time while the subjects were driving, as well as being given during post-test evaluations by rankings and ratings. The drivers were selected to represent the driving population and, particularly, to include those over age 55.

EXPERIMENTAL DESIGN, STATISTICAL ANALYSIS, AND DESIGN INTERPRETATION

The discussant's initial allegation suggests that there were errors in experimental design, statistical analysis, and data interpretation. Certainly, we have reputations for strength in the areas of rigorous design and statistical analysis. Because of this, we went to extra lengths, which included the use of advanced data collection technology, to have a most powerful experimental design.

Most researchers conducting field studies on pavement markings have not been fortunate enough to be able to incorporate the most powerful statistical design for identifying significant differences in pavement marking treatments. This most powerful design, which we used, is a repeated measures design with control. In general, most other studies do not collect data in a way that enables a given vehicle to be traced throughout the pavement marking zone. Therefore there is generally an inflated estimate of speed variability, and the resulting test statistic for determining the pavement marking treatment effects is not sensitive to small differences. By tracing the vehicles through the study site, we were able to incorporate the covariance structure among vehicles into the design (and test statistic) to produce a statistical test that is most powerful in declaring statistical significance among very small differences. The fact remains that even by using the power statistical method in the study, we found no statistically significant differences among the 4-, 2-, and 1-ft stripes on 40-ft centers with respect to speed, lateral placement, lane encroachments, and erratic maneuvers. In summary, the statistical design and analysis that we used are beyond reproach.

TEST SITES

Our work was a valid comprehensive field study representative of a variety of two-lane, two-way highways within the scope and budget available. In an attempt to generalize the results as much as possible, the studies were conducted in four states, with test sites that represented as diverse geometrics and characteristics as could be found on overlay projects on two-lane, two-way rural highways. We specifically stated the study conditions and limitations, and we suggested future research to resolve the issue of *optimum* centerline striping patterns for other kinds of work zone applications.

CONTRAST BETWEEN PAVEMENT AND MARKINGS

The discussant made an issue of the "extremely bright optical targets" (the stripes on black asphalt), claiming that they were easier to see than markings in most driving situations. Overlays on two-lane, two-way rural highways are, by definition, fresh, dark backgrounds. We would have been remiss to have used sun-baked irregular or worn surfaces because they do *not* represent the real-world situation with *fresh overlay*. The fact that the temporary tape markings provide excellent contrast and visibility is a point in their favor.

In the real world, temporary centerline stripes on overlay projects are in place for a period of up to 2 weeks until the overlay work is completed. A striping crew then applies permanent striping. Also, because the temporary yellow reflective tape markings used as centerlines for these projects are in place for a maximum of approximately 2 weeks, the stripes are indeed brilliant. The temporary centerline markings at pavement overlay projects on two-lane, two-way rural highways generally *are* superior in terms of cleanliness and reflectivity to the markings on upstream and downstream highway sections because they are newer. This is precisely one of the reasons that driver performance was the same when the 2- and 1-ft striping patterns were used, compared to the 4-ft striping pattern on the overlay sections of the rural highways. If the markings were in place for significantly longer durations, as is the case for the other work zone applications, then it is possible that the findings would be different.

EDGELINES

The discussant criticizes NCHRP for requiring no edgelines in the field study. Obviously, if long, continuous edgelines had marked the pavement course, the drivers might well have used these markings for visual guidance rather than the centerline stripes that were the subject of the research that we reported, and thus it would not have been possible to evaluate the specific effects of the three candidate striping patterns.

A second point is that edgelines would not be representative of many overlay projects on two-lane, two-way rural highways, where often the only cue the driver has is the centerline stripes. So, had we done the research with edgelines, the findings would be inapplicable to the more common overlay situation immediately after the pavement is laid.

Great care was taken to ensure that the administrative personnel and measuring devices would not be seen before the test site and thereby would not bias the drivers. This point was implied by the statement that these personnel were in a vehicle far from operating lanes.

DATA INTERPRETATION AND SAFETY IMPLICATIONS

It is in the area of data interpretation and safety implications that the discussant's flawed analysis is most evident.

Application of Results

We clearly emphasize and state that *the results only apply to pavement overlay projects on two-lane, two-way rural projects* and certainly do not advocate translating these findings to other highway work zone situations. Early in her comments the discussant falsely gives the readers the impression that we recommend adoption of the 2- or 1-ft striping pattern in *all* work zones. She admits near the end of her discussion that our paper states "that the findings should not be generalized to situations not tested."

Drivers' Evaluations

The results of the driver evaluations showed that on the average, although the 1-ft striping pattern was rated and ranked slightly lower *numerically*, its ratings and rankings were not statistically significantly different from those of the 2- and 4-ft striping patterns. Drivers at seven different sites *could have* rated the 4-ft striping pattern consistently the highest, but they did not. They were aware that the 4-ft stripe was longer than the 1-ft stripe but had some difficulty discriminating between the 2-ft and 4-ft stripes. They had trouble discriminating brightness as well. All patterns were judged "satisfactory" to "excellent." The discussant interpreted these ratings as less than desirable and suggested that the drivers were trying to indicate they wanted still longer stripes. It is true that in earlier research by TTI in *controlled* proving ground studies (not field studies), drivers *rated* the 8-ft stripe as their first choice. However, driver *performance* with the 8-ft stripe was not better than with several shorter stripes. Furthermore, the 8-ft stripe was not one of the treatments investigated in the field studies, as noted above. Only if the drivers had *consistently* rated the 4-ft stripe as the best would a still longer stripe have been indicated in the present application. Further, seldom does a sample of drivers rate *anything* as "excellent." Driver variability enters the picture in rating, as does a tendency to include other environmental elements into the ratings (weather, previous highways driven, etc.).

Cost

The criticism that the study narrowly interprets the word "cost" when lives are at stake was an attempt to discredit the research as leading to a loss of lives. Only if someone grossly misinterpreted the objective of the study would this issue apply.

Presentation of Results

The discussant implies that because the research that we reported did not evaluate all possible combinations, it should not have been reported at the Annual Meeting of the Transportation Research Board. If the research community waited until all variables relative to a subject are evaluated before presentation and publication, knowledge would not be advanced very much.

ACCEPTANCE OF RESULTS

The discussant's review contradicts previous reviews, which are based on the knowledge and integrity of reputable researchers and highway officials. The study results were reviewed and accepted by these knowledgeable professionals: (a) an NCHRP Panel of Experts, (b) the TRB Committee on Traffic Control Devices, and (c) experts serving on the Construction and Maintenance (C&M) Technical Committee of the National Committee on Uniform Traffic Control Devices. On the basis of the results of the research and recommendations from a task force headed by one of us, the C&M Technical Committee unanimously approved a recommendation that a full complement of markings be used in all work zones, with the exceptions noted in the following recommendations:

> The National Committee requests that the FHWA reopen and revise the Rule-Making on Work Zone Pavement Markings which appropriately reflects the following recommendations:
>
> 1. For paving operations, short-term markings may be installed to a lesser dimensional standard than that specified for permanent markings.
>
>   Short-term pavement markings for paving operations are defined as temporary pavement marking lines placed on centerlines and lane lines, following the paving operations, which will be in place up to two weeks, at which time it is expected that permanent markings will be in place.
>
>   To the extent practicable, it is intended that temporary work zone markings and/or appropriate channelizing and delineating devices will approximate the guidance normally supplied by permanent markings.
> 2. Short-term pavement markings for lane lines and dashed centerlines may be less than four feet in length.
> 3. The National Committee recommends that the FHWA recognize that normal pavement markings for chip and sand seals on low-volume roadways, and roadways undergoing milling operations, may not be practical and therefore other delineation treatments shall be used.
> 4. When the installation of short-term pavement markings is impractical during pavement operations, channelizing or delineating devices with appropriate warning signs shall be used.
> 5. The National Committee endorses additional research to be conducted by FHWA to improve engineering practices to insure that safe and cost-effective temporary markings and other delineation treatments are adopted for use in highway work zones, particularly for short-term paving operations.

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