TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
SWUTC/95/721917-1		

4. Title and Subtitle	5. Report Date
Field and Laboratory Studies of Warning Symbol Sign Legibility Distance	March 1995
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Frances A. Greene, with contributions by Rodger J. Koppa, Ronald D. Zellner, R.Dale Huchingson, J.J. Congleton and A. Garcia-Diaz	RF 721917-1
9. Performing Organization Name and Address	10. Work Unit No.
Texas Transportation Institute Texas A&M University System College Station, TX 77843-3135	11. Contract or Grant No. DTOS 88-G-0006
12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
Southwest Region University Transportation Center Texas Transportation Institute The Texas A&M University System College Station, TX 77843-3135	Study Report (1 of 2)
	14. Sponsoring Agency Code

15. Supplementary Notes

Supported by a grant from the U.S. Department of Transportation
University Transportation Centers Program

16. Abstract

Laboratory studies of warning symbol signs have been shown to underestimate legibility distances by up to a factor of two when compared with field studies. However, the research reported here suggests it may be more than just experimental settings contributing to disparity in research findings. For this research a new laboratory simulation technique was developed. The methodology optimized factors criticized in earlier studies, thus increasing fidelity. Correlation coefficients between laboratory and field legibility distances were very promising. The newly developed laboratory simulation was a successful first step at correcting problems associated with laboratory studies in the past. It is argued that recommended distances at which signs are to be placed, especially with consideration for older drivers, should be determined from a "worst-case" scenario of the minimum distance with tolerances.

Large within-subject variability, evident in both young and older drive age groups, overshadowed contributions of experimental variables. No studies have reported this finding of large within-subject variability in similar research. This variability led to an analysis of variation in distance estimation if different strategies for numbers of data points collected are used.

17. Key Words	18. Distribution Statement
Warning signs, legibility distance, older drivers, laboratory study, field study, within subject variability	

19. Security Classification (of this report)	20. Security Classification (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	156	

Form DOT F 1700.7 (8-69)

FIELD AND LABORATORY STUDIES OF WARNING SYMBOL SIGN LEGIBILITY DISTANCE

A Research Study
by
Frances A. Greene

With Contributions by
Rodger J. Koppa
Ronald D. Zellner
R. Dale Huchingson
Jerome J. Congleton
Alberto Garcia-Diaz

Southwest Region University Transportation Center
Texas Transportation Institute
The Texas A&M University System
College Station

March 1995

DISCLAIMER

ACKNOWLEDGEMENT

The authors would like to acknowledge the partial support provided by a grant from the U.S. Department of Transportation, University Transportation Centers Program to the Southwest Region University Transportation Center, which made this research possible. The remainder of the funding was provided by the Department of Industrial Engineering, College of Engineering, Texas A&M University.

EXECUTIVE SUMMARY

ABSTRACT

Laboratory studies of warning symbol signs have been shown to underestimate legibility distances by up to a factor of two when compared with field studies. However, this research suggests it is more than simply experimental setting contributing to disparity in research findings. Using a group of old and young drivers, six symbol signs were investigated in both settings. With six trials per sign, legibility distances, defined as the distance at which the sign is correctly identified from a menu, were collected.

Two major findings emerged from this research. The first is the successful development of a new laboratory simulation technique for the investigation of legibility distances. This new laboratory simulation technique was designed as a methodology which optimized factors criticized in earlier studies, thus increasing fidelity. Correlation coefficients between laboratory and field legibility distances were computed and appear very promising. The newly developed laboratory simulation was a successful first step in correcting problems associated with laboratory studies of the past.

Large within subject variability was discovered in both age groups' data of legibility distances, both in the laboratory and field trials.. This variability led to alternative ways of defining the dependent variable equivalent to designs of past studies examining legibility distances of the same signs. Different results arose out of the subsets created. The consideration is not just should a field-based versus laboratory-based methodology be used. An argument is posed that recommended distances at which signs are placed must be determined from a "worst-case" scenario. This premise requires a reexamination of our research methodologies for determining placement of highway signs.

INTRODUCTION

Research on traffic sign legibility spans over five decades. A dichotomy exists in paradigms used in this research; on one hand, laboratory-based studies and on the other, field studies. For many years researchers have discussed and debated the merits and problems associated with methods chosen for field and laboratory studies to determine legibility distances for highway signs. Dewar (1973) reviews the categories of methods used for traffic sign research. Dewar (1973) performs an excellent critique of both laboratory and field studies' methodologies; many of the criticisms still valid two decades later. Recently, Zwahlen, Hu, Sunkara and Duffus (1991), report discrepancies found between legibility distances from laboratory versus field studies for the same signs. They compare legibility distances for warning symbol signs found by different researchers for identical symbol signs. When Zwahlen et al. (1991) compared legibility distances for warning symbol signs found in previous research, they found for laboratory studies the same signs have shorter legibility distances than those reported from field studies. Zwahlen et al. report their field data indicate daytime field legibility distances to be double those of laboratory distances. A plethora of hypotheses for the underestimation of legibility distance obtained in laboratory studies are given, including several involving display devices chosen for sign stimuli presentation. A few cited are insufficiency of display device fidelity, luminances and contrasts not controlled or specified, insufficient display resolution, small image vibrations on monitors or projectors, non-uniform and insufficiently sharp symbol contours. Zwahlen et al. (1991) speculate laboratory legibility distance results could be a product of display devices chosen for stimuli presentation which make the laboratory setting unlike the real world and introduces extraneous variables.

Zwahlen et al.'s (1991) research poses an interesting question. Which studies are reporting the "best" legibility distances for warning symbol signs? Or stated another way, what is the "correct" methodology to choose when collecting data to be used for traffic signs' placement guidance? A review of the literature does not find one study which evaluates the relative effectiveness of the two methodologies, field and laboratory-based, in order to make a clear-cut choice of which to choose. A case can be made that field-generated data is preferable, as it has maximum face validity for application of results. Is the decision to use a field-based versus laboratory-based methodology the only question a researcher must consider? Taking this assertion one step further, the present research asks the question: disregarding whether a field or laboratory study is performed, are experimental design and analyses of data without regard to within subject variability inherent in a driver having a dramatic impact on the reported results? It will be argued that recommended distances at which signs are placed to be recognized should be determined from a "worst-case" scenario. This supposition requires a reexamination of our research methodologies for determining guidelines for placement of symbol signs.

Zwahlen et al.'s (1991) conclusions about laboratory-based legibility research's questionable real-world validity and challenges to identify/control factors responsible for smaller recognition distances were the influences for this research. Optimization of and control for as many shortcomings of previous laboratory methods as Zwahlen at al. (1991) iterate were objectives behind developing a new laboratory simulation technique. This study consisted of a field study and a laboratory study, using that simulation technique. Legibility distances for the same warning symbol signs with a sample of drivers widely varying in age were collected in both settings.

This study applies different analyses of the legibility distances collected from young and old drivers in both daytime field and laboratory studies. The research is aimed at exploring not only the experimental setting, but also implications for the methodology used to collect data. Traffic warning symbol signs are increasing in popularity and usage on American highways due to assets of ability to communicate a warning message and international understanding. With over 50 years of research being devoted to highway signing, two kinds of research paradigms have emerged: laboratory-based and field studies. Merits as well as deficiencies associated with the two methods to determine legibility distance for highway signs have been debated. Dewar (1973) reports a major problem with laboratory studies is an inconsistency with replicating conditions of a driving task, inasmuch as there are no normal visual cues nor distractions of attention; the participant is merely a passive observer. Field studies, on the other hand, ignore drivers' experience with signs and may not use representative samples of drivers.

FIELD STUDY

Method

Subjects. Twenty-four subjects served as participants, equally divided between young (mean age 23.7 years) and older (mean age 73.2 years) drivers. Visual acuity (corrected) for the older subjects ranged from 20/20 to 20/50 and from 20/13 to 20/40 for the young drivers. An equal number of males and females constituted the young group, while 7 males and 5 females composed the older group.

Procedure. A training session was conducted where subjects were familiarized with a 27-symbol sign menu until they reached a 100% identification criterion. The drivers used the menu at all times and were instructed to respond by sign number only when they were certain as to the sign identification. The certainty aspect of their response was stressed and was equated willingness to make a vehicle maneuver in response to the sign. Subjects drove an instrumented vehicle with distance-measuring equipment toward a warning symbol sign at a speed of 40 km/h on a 610m straight taxiway section of Riverside Campus, Texas A&M University. The distance readout from the equipment, plus the subjects' verbal sign response were recorded on video tape.

Stimuli. Nineteen new, 9.45cm by 9.45cm warning symbol signs served as stimuli. Of the 19 signs, only six were analyzed for legibility distances; the other 13 served as distractors. The stimuli were exactly the same for both the field and laboratory studies. The six target signs and their *Manual on Uniform Traffic Control Devices* (1988) designations are in Table 1.

TABLE 1

Symbol Signs Used in Field and Laboratory Studies

Symbol Sign	MUTCD Designation
Crossroad	W2-1
T-Junction	W2-4
Two-Way Traffic	W6-3
Slippery When Wet	W8-5
Bicycle Crossing	W11-1
Deer Crossing	W11-3

Six responses for each of the six target signs shown in Table 1 were collected from each driver. Sixty trials, a random selection of target and distractor signs, were run. All data were gathered during daytime conditions, with signs northerly facing to prevent glare from the sign face. Subjects participated in the field study first, with the laboratory study starting nine weeks later.

Independent Variable. Gender and age group (young or older) were between subject variables. Within subject variables were individual sign and condition (laboratory or field). A repeated measures design was chosen.

Dependent Variable. Legibility distance, defined as the distance from the warning sign where the driver applied the "certain" criterion to its identification, was the dependent variable measured.

LABORATORY STUDY

Subjects, target and distractor signs, number of replications and trials in the laboratory study were identical to the field study. Independent and dependent variables remained the same.

Experimental Apparatus. A Macintosh IIci computer and a Viewsonic 43.2cm high-resolution, color monitor were used to present the newly developed simulation of the field study. Graphics (Deneba Canvas™ 3.0) and animation software (MacroMind Director® 3.1.1) were used to create and present the simulation. Hypercard© 2.1 controlled the experimental session, collected and stored all responses.

Simulation Development. Animation was a key element of the new simulation technique. A background scene from the field study was videotaped and digitized at 24-bit resolution. A graphics program was used to draw images of the signs in 32-bit resolution color and scale to correct percentage for size, to replicate retinal image size in 3.05m increments from 610m to 24m. The resulting 184 graphic sign images were embedded one at a time in the digitized scene background. The animation program, with the sign elements placed in the background, created a simulation of approaching a sign at a speed of 33 km/h. Subjects were seated 6.1m from the monitor.

Procedure. Using the identical training standard and procedure as the field study, the equivalent "certain" criterion for legibility distance response was applied. The experimenter input subjects' responses via keyboard and the computer verified the correctness of the sign response and recorded legibility distance for all 60 trials.

RESULTS

Legibility Distance

Consistently, older drivers had a shorter legibility distance for all signs than young drivers. For both the field ($F(1,20) = 10.38$, $p<0.004$) and laboratory ($F(1,20) = 10.51$, $p<0.004$) conditions, the main effect of age was significant. In the field, older drivers' mean legibility distance was 68% that of young drivers, with young drivers' legibility distance at 342.11 meters, while older drivers' overall mean legibility distance was 232.64 meters. Laboratory data showed the same trend, with the mean legibility distance for older drivers being 263.6 meters (864.7 feet); young subjects had a mean legibility distance of 372.2 meters (1221.3 feet), or 71% of the legibility distance of young drivers.

The main effect of individual sign was also significant for both field and laboratory studies. The rank order of signs by legibility distance was the same for field and laboratory studies. Figure 1 plots mean legibility distance by sign for both conditions. As Figure 1 displays, legibility distance for all signs was longer in the laboratory than the field, in contradiction to Zwahlen et al.'s (1991) findings.

For both the field and laboratory studies, the main effect of sign was significant $F(5,100) = 70.3$, $p<0.0000$ for the field study and $F(5,100) = 54.2$, $p<0.0000$. Table 2 summarizes the mean field and laboratory legibility distances for the six signs in increasing order of legibility distance, collapsed over age.

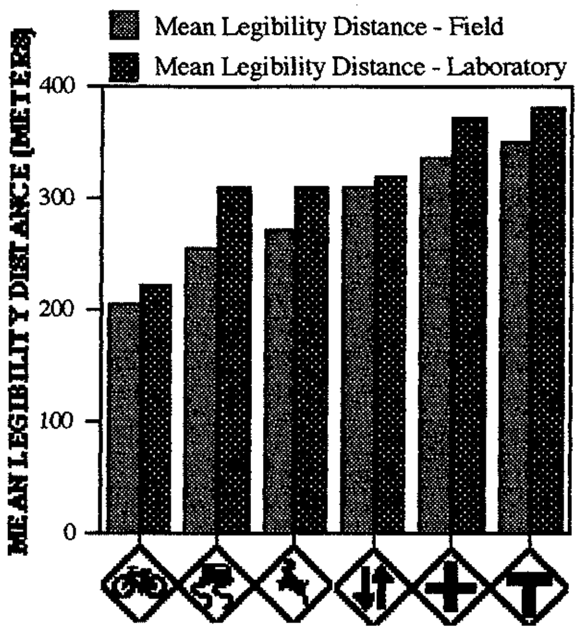

Figure 1. *Legibility distance by sign and condition*

TABLE 2

Mean Legibility Distance for Six Signs, Field and Laboratory Data

Warning Sign	Field Mean Legibility Distance	Laboratory Mean Legibility Distance
Bicycle Crossing - W11-1	205.01 m	220.45 m
Slippery When Wet -W8-5	254.35 m	308.27 m
Deer Crossing - W11-3	270.20 m	308.82 m
Two-Way Traffic - W6-3	308.27 m	317.57 m
Crossroad - W2-1	336.15 m	371.41 m
T-Junction - W2-4	350.25 m	381.00 m

Comparisons Between Field and Laboratory Legibility Distances

One of the original rationale for the design of this research was to develop a new laboratory simulation technique whose fidelity approached that of the real-world. In order to test if the laboratory legibility distances could be used to predict field legibility distances for new traffic signs, some measure of the correlation between the mean legibility distances for the six signs was necessary. A Pearson product-moment correlation was calculated for each warning symbol sign. The y_i is the field mean legibility distance. The y_j is the average legibility distance response collected in the laboratory portion. The computed correlation coefficient, r_{ij}, for each of the six target warning signs under study is presented in Table 2.

TABLE 2

Correlation Coefficients for Mean Legibility Distance, Field and Laboratory Data

Warning Sign	**Correlation Coefficient, r_{ij}**
T-Junction - W2-4	0.8359
Crossroad - W2-1	0.8340
Two-Way Traffic - W6-3	0.8195
Deer Crossing - W11-3	0.7822
Slippery When Wet - W8-5	0.7711
Bicycle Crossing - W11-1	0.6484

The larger r_{ij} values for predicting laboratory legibility distance from field legibility distance are encouraging. The value of r_{ij} expresses the degree of relationship between the two response variables: field legibility distance and laboratory legibility distance for the six warning symbol signs. The larger values of r_{ij} for the warning symbol signs *Two-Way Traffic*, *Crossroad*, and *T-Junction*, indicate that the laboratory data is a good predictor of field legibility distance. However, that statement is not true for *Bicycle Crossing*, as it has a lower value of r_{ij}, too low perhaps for a traffic engineer to accept the prediction.

Within Subject Variability

Individual plots of legibility distance for the six replications for each sign for each subject revealed large within subject variability. Due to the large amount of within subject variability among the six responses for both the field and laboratory conditions, a coefficient of variation (CV, computed as mean/standard deviation * 100) was computed for each condition, age group and sign.

The CVs ranged from values 2.04 to 49.17 for the field condition and 2.8 to 41.4 for the laboratory condition. Older subjects were more variable in their responses than the younger group (*Mean* = 17.25 for older subjects; *M* ean = 12.53 for the young group). The *Crossroad* sign had the least amount of variability (*M* ean = 9.85); while *Bicycle Crossing* had the most (*M* ean = 20.39).

This large within subject variability suggests that using the mean of the six replications of legibility distance for each sign could lead to misleading results and drawing erroneous conclusions. Therefore, additional analyses of the six replications were examined. Analyses included the following subsets of the original legibility distances:

a. First two responses
b. Minimum of the first two responses
c. Overall minimum response

Of course, all six replications were also included in the analyses.

Figures 2 and 3 show the legibility distance obtained for each of dependent variables described above for two signs: *Bicycle Crossing* and *Cross-road.* Figure 2 summarizes the data for older subjects only, and Figure 3 for young subjects only. The two signs chosen for analyses represented extremes in variability, with *Bicycle Crossing* having the most in both field and laboratory studies.

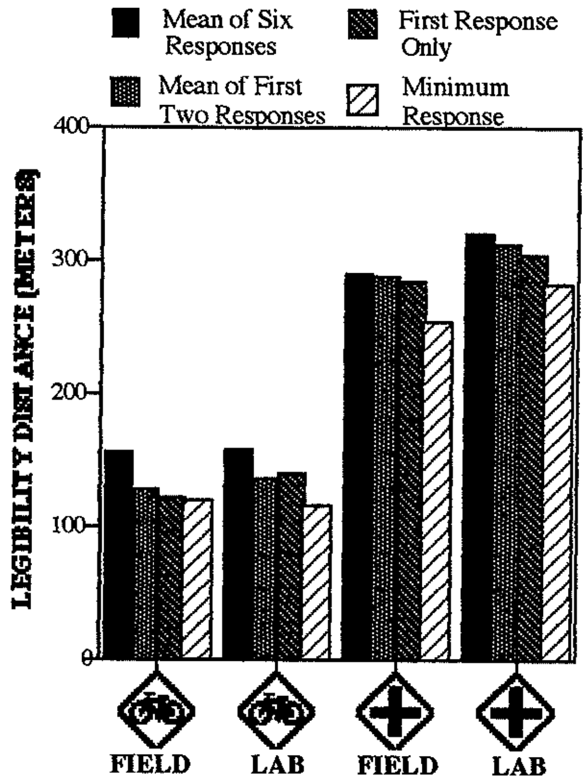

Figure 2. *Older drivers only, subsets of data responses, both conditions*

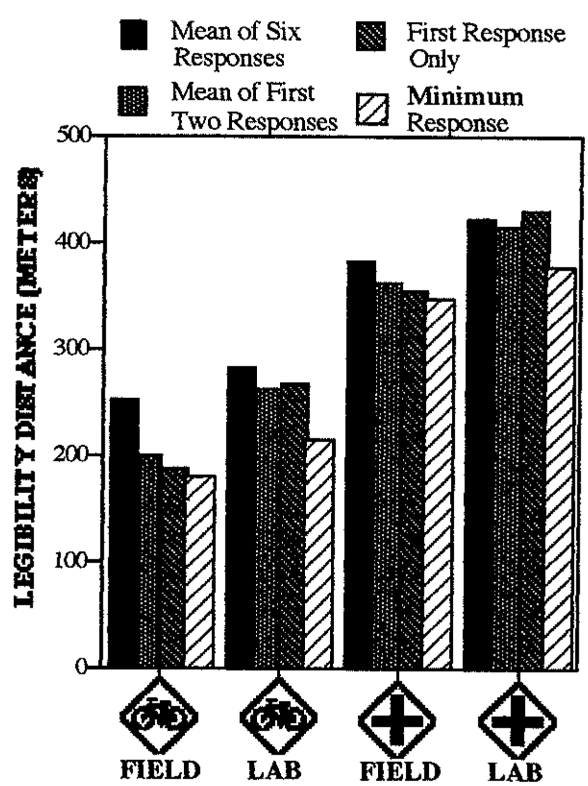

Figure 3. *Young drivers only, subsets of data responses, both conditions*

DISCUSSION/CONCLUSIONS

With regards to the findings of large within subject variability for both age groups, in both settings, the data depicted in Figure 2 (older drivers only) dramatize the differences in conclusions which would be drawn for legibility distances of the six warning symbol signs. The data in the four columns for each warning sign represent determinations one would make for recommended legibility distances for the six signs in the following manner:

Column 1 - Mean of all six responses - the conclusions drawn from this study
Column 2 - First response only - equivalent to Paniati (1988) and any study which included no replications
Column 3 - Mean of first two responses - Zwahlen et al. (1991), since he collected two data points on each of his young subjects.
Column 4 - Worst-case scenario. This minimum of six responses from this study is the shortest legibility distance and represents an extremely conservative recommendation. These signs would be legible to a representative sample of older drivers (over the age of 65).

Figures 2 and 3 depict graphically that depending on the subset of dependent measures used in the analyses, the legibility distances reported are quite different. This difference is attributable to the large within subject variability.

Most studies either do not report the number of replications or do not include any. In Paniati (1988), only one data point is collected and the average legibility distance is reported for each sign as a function of subject age. Kline, Ghali, Kline and Brown (1990) also only collected one data point per person. Their design is comparable to reporting the "First response only" column of data from this study. Zwahlen et al. (1991) include two replications in their study. They also plot the mean, standard deviation, minimum and maximum for legibility distance for 12 signs investigated. However, no tabular data is reported and it is difficult to interpolate the data from their figure plot. Zwahlen et al.'s (1991) study results would equate to the reporting of the second column, "Mean of the first two responses " data from this study. By not examining the within subject variability and reporting the mean of the six responses for older drivers' field condition overestimates the legibility distance for *Bicycle Crossing* by almost 28% over the first response only and 22% over the mean of the first two responses. *Bicycle Crossing* was the sign with the most within subject variability.

In addition to answering the shortcomings of laboratory studies described by Zwahlen et al. (1991), this simulation technique was hoped to be the first step toward creation of a simulation tool for traffic engineers. This study was aimed at developing a prototype of that tool. Several interesting findings emerged from this research. First, the recognition of the large within subject variability has been discussed in an earlier paper (Greene et al., 1994). Mean legibility distance for all signs, regardless of age, was always longer in the laboratory study than in the field study - a result which contradicts Zwahlen et al.'s (1991) finding. Conceivably the fact that all subjects participated in the field study first caused a learning effect which could account for the longer laboratory legibility distances.

Overall, results of development and application of this simulation technique were very encouraging as a preliminary evaluation tool. This study was very possibly the first of its kind, using a repeated measures design, whereby identical young and older subjects participated in both a field and laboratory setting. The same warning symbol signs were investigated in both cases.

Another contribution this research has made to transportation literature is the successful development and proven use of a new laboratory simulation technique. This promise of this technique was demonstrated by the high correlation coefficients between laboratory and field legibility distances for all signs. Zwahlen et al.'s (1991) criticism of laboratory studies was valid. Laboratory studies conducted before this new technique was developed were lacking in real-world fidelity. The prototype simulation created for this research is a favorable initial step toward maximizing parameters which may have caused the dissimilarity in results between prior field and laboratory studies.

Given the high cost and other uncontrolled extraneous variable difficulties associated with field studies, the laboratory alternative technique developed and used in this study would be preferred, especially in view of the high positive correlations between the legibility distances collected in the field and the laboratory. Future modifications and enhancements are possible with today's ever-changing face of multimedia technology. Furthermore, higher resolution display devices with higher luminance capabilities are becoming an accessible, as well as affordable option.

RECOMMENDATIONS

It was the close examination of data leading to the discovery of tremendous within subject variability for all signs, age groups and both conditions, when questions arose about what is the appropriate subset of dependent variable to use. This question raises the issue of experimental design and offers a suggestion of designing for a worst-case scenario.

It has been shown repeatedly that older subjects have shorter legibility distances for symbol and text signs than their young driver counterparts (Allen, Parseghian and Van Valkenburgh, 1984; Evans and Ginsburg, 1985; Kline, Ghali, Kline and Brown, 1990; Lum, Roberts, DiMarco and Allen, 1983; Paniati, 1988; Sivak, Olson and Pastalan, 1981). This study shows the same result. Therefore, it follows that legibility distances should be determined using a representative sample of older drivers, not young drivers with often better than 20/20 visual acuity.

It is recommended that the area of within subject variability in legibility distance responses be given serious examination and recognition for its existence. Also, the ramifications this variance has on predictions must be reconciled.

Researchers must come to grips with several issues of foremost importance: first, a representative sample of older drivers must be included in every study. The literature supports time and time again problems the aging process has on the visual system and reaction time. To continue the practice of using only young subjects to generate recommendations for sign placement is not recommended and ignores an important and growing segment of our population, the older driver.

Second, replications are vital and must be included in every study, whether it is field- or laboratory-based. Without replications, we, as practicing transportation researchers, are fooling ourselves if we think we can collect just one or maybe two observations and make recommendations. The conclusions in the literature which are based on only one observation or perhaps two are very likely misleading and possibly in error. Without replications, the author is overlooking that large and consequential within subject variation contribution.

Third, authors must completely define their response variable. The literature is full of different independent variables like recognition, legibility, visibility and identification distance. Most are not adequately defined, or defined at all.

Replications appear to be very relevant to the results which are reported. The research community must first recognize the existence of and then focus on the source of this within subject variation phenomenon. Without a study of this factor, perhaps that which is being reported in the literature is more a function of the experimental design than the multitude of independent variables being painstakingly manipulated.

Finally, the newly developed laboratory simulation technique appears to have increased the real-world fidelity by optimizing factors criticized by Zwahlen et al. (1991). The fidelity may have increased so much that it caused the reversal in trend in laboratory to field legibility distances comparisons. Previous literature always found shorter distances in the laboratory when compared to the field. However, this study produced longer legibility distances and certainly ones which had a reasonably high correlation with field distances. The simulation technique developed here was by no means the best available. But it used low-cost, off-the-shelf hardware and software. Future simulations can take advantage of the larger, higher resolution monitors, with more colors available at an affordable price. Other software programs are marketed on a consistently frequent basis. This new software may cut down the development time and enhance the fidelity even more. Future research should focus on low-cost hardware, with as many real-world features in the software development as possible. This does not mean having to procure a driving simulator. The technique used here had encouraging correlation coefficients when comparing laboratory to field data. Where the correlation coefficients fell short, additional work should be done to examine and correct the factors which led to the lower correlation coefficients.

TABLE OF CONTENTS

Page

Page

LIST OF FIGURES

Figure Page

Figure | Page

LIST OF TABLES

Table Page

Table Page

INTRODUCTION

BACKGROUND

Research of traffic sign legibility spans over five decades. With over 50 years of work being devoted to highway sign systems, one might assume all design and placement questions have been answered satisfactorily. In actuality, symbolic signs, which are reasonably new but becoming more popular in usage by most states, possess many parameters which are in need of further research. For textual signs, like the one in Figure 1, some of the factors cited most often as influencing legibility are:

a. letter height
b. stroke width of letters
c. spacing between letters and words
d. font type
e. capital or lower case letters
f. ratio of message size to total area of sign
g. contrast between figure and background

Figure 2 illustrates the same sign in symbolic form. A review of legibility factors above reveals that only the last two are relevant to symbol signs. The *Deer Crossing*

Figure 1. *Textual form of Deer Crossing sign*

sign is one of the most complex symbolic signs and does not lend itself to analysis such as optimum letter design.

Jacobs, Johnston and Cole (1975) describe how much more difficult it is to predict legibility of symbol signs than textual ones. They note component details of symbols

Figure 2. *Symbolic form of Deer Crossing sign*

vary in size and with the exception of simple signs such as *T-Junction* and *Crossroad*, the elements analogous to stroke width of a letter are impossible to quantify. Markowitz, Dietrich, Lees and Farman (1968) believe symbols should communicate their meaning without unnecessary details.

Dewar (1973) emphasizes a symbol sign should not be too visually complex or bear similarity to other signs, so as to be confused. Bowen (1993) hypothesizes that legibility of symbols is related to the spatial frequency content of the sign. Bryant (1982) notes design of symbols is more an art than a science and concedes that it has not received much attention. Hence, Bryant, recognizing shortfalls in symbol sign design research, calls for comprehensive field testing of symbol signs. Studies investigating physical requirements of text signs to enhance their legibility do not apply directly to the design of symbol signs. Therefore, an increase in research efforts of legibility factors for symbol signs is needed and being seen in the literature. There are two kinds of paradigms used in this research: laboratory-based studies and field studies.

For many years researchers have discussed and debated the merits and problems associated with methods chosen for field and laboratory studies to determine legibility distances for highway signs. Dewar reviews categories of data collection methods, as well as dependent variables measured in traffic sign research. Dewar performed a thorough critique of both laboratory and field studies' methodologies; many of the criticisms continue to be valid two decades later. He considers a major problem with laboratory studies to be their inconsistency with replicating conditions of a driving task, i.e., there are no normal visual cues nor any distractions of attention - the participant is merely a passive observer. Field studies ignore drivers' experience with signs and may not use representative samples of drivers. Dewar outlined the multitude of methodologies utilized to investigate highway sign legibility. He asks, "With several methods available to measure each of a number of variables one might ask, 'Which method is best for my particular need?'" He does not feel as though the relative effectiveness of the two methodologies has been adequately evaluated.

When results of legibility distance research for specific symbol signs are compared, a wide variability in reported distances is found. Recently, Zwahlen, Hu, Sunkara and Duffus (1991), report discrepancies they uncovered between legibility distance results from laboratory versus field studies for identical symbol signs. Zwahlen et al. (1991) compared the legibility distances for warning symbol signs found by different researchers. The authors find that for studies conducted in a laboratory, the same signs have shorter legibility distances than those reported from field studies. Table 1 summarizes those differences in legibility distance by symbol sign for four different researchers.

TABLE 1

Summary of reported daytime mean legibility distances for selected symbol signs

Study/Type	**Mean Legibility Distance: Crossroad - (W2-1)***	**Mean Legibility Distance: Bicycle Crossing - (W11-1)***
Paniati (1988) /Laboratory Study	215 m (705)	109 m (358)
Jacobs, Johnston, and Cole (1975)/Laboratory Study	500 m (1640)	Not Investigated
Zwahlen, Hu, Sunkara, and Duffus (1991) /Field Study	428.2 m approximately (1405)	213.4 m approximately (699)
Bowen (1993) /Laboratory Study	633.9 m (2080)	328.4 m (1076)

* **numbers in parentheses denote distances in feet**

If one were attempting to predict legibility distance for any of the three signs compared in Table 1 above, the differences in results are so enormous it would be impossible to reach any agreement on a choice. As can be seen from Table 1, for the *Crossroad* sign, Paniati (1988) and Bowen (1993) differ in their distances by a factor of almost 3:1. For *Bicycle Crossing*, comparing legibility distances for Paniati (1988) and Zwahlen et al. (1991), the difference is by a factor of 2:1; while the difference obtained by Bowen is 3:1. In addition to the inconsistencies Zwahlen et al. (1991) point out, the differences in results are larger between laboratory studies than between field and laboratory studies.

Past Research on Symbol Signs

Examining research studies which investigated daytime legibility distance for symbol signs, the following is a summary of the relevant aspects of experimental design and response variable, as related to the present research.

Jacobs, Johnston and Cole (1975), used both symbol and text signs. However, since their signs met the guidelines of the *Australian Road Signs Code* (1960), only the Australian *Crossroad* sign was of similar enough design to be compared with the *Manual on Uniform Traffic Control Devices* (MUTCD), 1988, designation W2-1, *Crossroad* sign used in other studies. Ten subjects participated and no ages were given, although the subject group had a mean visual acuity of 6/4.4 (or approximately 20/13), therefore, it is fairly certain that they were not all older drivers. Only black and white slides were used and the response was stating were not all older drivers. Only black and white slides were used and the response was stating the meaning of the sign. The authors do not report if replications were used. The responses were judged correct or incorrect and 50 and 95 percent threshold legibility distances were calculated. The 50 percent threshold legibility distance for *Crossroad* was reported to be approximately 500 m.

Paniati (1988) investigated legibility distances for 22 symbol signs using 32 subjects, half young (mean age 33) and half old (mean age 61). Color slides of signs were presented varying visual angles equivalent to that of full-sized signs at a distance. The dependent variable was legibility distance, as determined when the symbol was "identified" and the features of the symbol described. It is assumed each subject saw each sign only once. Of the 22 signs investigated, Paniati publishes legibility distances for the same six signs used in the present research. Those distances are accounted in Table 2.

TABLE 2

Legibility distances from Paniati (1988) for six symbol signs

Sign and MUTCD Designation	Mean Legibility Distance Overall (meters)	Mean Legibility Distance - Young Group (meters)	Mean Legibility Distance - Old Group (meters)
Crossroad (W2-1)	215	248	181
T-Junction (W2-4)	178	228	128
Two-Way Traffic (W6-3)	142	150	134
Bicycle Crossing (W11-1)	109	148	78
Slippery When Wet (W8-5)	66	76	55
Deer Crossing (W11-3)	49	54	44

Kline, Ghali, Kline and Brown (1990) reported visibility distances for young, middle-aged and older drivers for text and symbol signs under day and dusk conditions. Sixteen subjects in each of the three age groups (young-$M = 24.2$; middle-aged-$M = 53.7$; elderly-$M = 66.5$) viewed four symbol signs: *Divided Highway*, *Road Narrows*, *Men Working* and *Hill*. The signs were presented on a monochrome computer monitor, therefore, color of the warning sign was not a factor. In a similar method as used by Paniati, the signs started out too small to be recognized and were slowly increased in size until the sign could be correctly described using a "blind-person criterion." This response was a description of the sign that would be sufficient enough to convey the content to someone who could not see it. The smallest size at which the sign was accurately described was the threshold for that sign. The dependent variable, visibility distance, was calculated using the equation :

Visibility Distance = [(Sign Height ÷ Threshold Height) x Viewing Distance]

Again, it is assumed each subject saw each sign only once. Kline et al. present their data in such a manner as to make it impossible to determine visibility distances for individual signs. It must be noted here that each author often has a unique definition of their dependent variable(s). Kline et al. used *visibility distance*, which is related to sign description at a threshold level. Their *visibility distance* is not defined in the same manner as *legibility distance* for this research. Kline et al. use a different response variable, which is simply another means of approaching sign legibility.

Zwahlen, et al. (1991) summarize their findings of field "recognition distances" for 10 warning symbol signs. Only young subjects (mean age = 29.8 years) participated in a daytime field study. Responses consisted of the subject notifying the experimenter as soon as he/she "recognized" the symbol on the sign. That distance from the sign composed the dependent variable. Zwahlen et al. never define the criteria for sign recognition. Each sign was seen twice by each subject. Zwahlen, et al. plot mean, ± one standard deviation and minimum and maximum legibility distances. However, due to the manner in which the table of results is given by Zwahlen, et al., it is difficult to tabulate exact recognition distances. However, an approximation for the five signs of concern is presented in Table 3.

TABLE 3

Reported recognition distance from Zwahlen et al. (1991) study

Sign and MUTCD Designation	**Average Recognition Distance in meters (Approximate)**
Crossroad (W2-1)	428.2
Two-Way Traffic (W6-3)	373.4
T-Junction (W2-4)	358.1
Slippery When Wet (W8-5)	221.0
Bicycle Crossing (W11-1)	213.4

Bowen (1993) used a laboratory study to generate recognition distances for *Crossroad, T-Junction, Slippery When Wet, Two-Way Traffic, Bicycle Crossing* and *Deer Crossing* signs. Fifteen subjects, only two of age ≥ 55 years, participated. The procedure involved cards containing miniature high contrast black and white signs moving toward the participant until he/she made a correct match between the stimulus sign and one on a menu available at all times. The participants matched their perception of the target sign with the a sign on the menu by identifying it by number. A correct match resulted in a recognition distance for each symbol sign. No naming of signs was required, nor interpretation of sign meaning or representation.

The results from Bowen's study are summarized in Table 4. This table includes data on the coefficient of variation (CoV) from his study. This measure of relative dispersion, expressed as a percentage, is defined as:

$$\text{CoV} = s \;/\; \overline{X}, \text{ where } s \text{ is the standard deviation and } \overline{X} \text{ is the mean.}$$

CoV is a measure of relative dispersion since it is measured with respect to the mean.

TABLE 4

Mean recognition distance and Coefficient of Variation from Bowen (1993) for six warning symbol signs

Sign and MUTCD Designation	Recognition Distance (meters)	Coefficient of Variation (CoV)
Crossroad (W2-1)	634.0	0.359
T-Junction (W2-4)	643.1	0.339
Two-Way Traffic (W6-3)	418.2	0.283
Slippery When Wet (W8-5)	403.6	0.282
Deer Crossing (W11-3)	499.6	0.424
Bicycle Crossing (W11-1)	328.4	0.442

Of the studies aimed at determining legibility distances for warning symbol signs, Bowen (1993) is the only author who used more than two replications in his repeated measures design. Each participant viewed the six signs four times during the trials. These data collected allowed more scrutiny of within subject variability. Bowen (1993) writes "The range of variation about the mean for each sign is seen to be large, and the distributions do not appear to be normal. ..Relative dispersion of recognition distance measurements for each sign is large, as indicated by its coefficient of variation (CV), defined as the standard deviation divided by mean value." Tabular values of coefficient of variation are also contained in Table 4. It can be seen that the signs with the largest relative dispersion are *Bicycle Crossing* and *Deer Crossing*, with the standard deviation being over 44% and 42% of the mean, respectively.

Bowen also points out that his data contained a small, but statistically significant, increase in recognition distance over the four trials. This indicates some form of learning or familiarity from the exposure to the same signs is occurring over repeated trials. Since the recognition task is actually one of matching a perception with that same pattern on the menu, a process of elimination is possibly resulting as trials continue. Bowen postulates it is easier to see what sign it is not and narrow the possibilities down from an absolute number of 27 to a much smaller subset which meet initial parameters of similar shape, orientation, etc.

As described above by the review of the consequential studies which have researched legibility distances for warning symbol signs, in addition to what Zwahlen et al. (1991) point out, the differences in results are larger between laboratory studies than between field and laboratory studies. However, a direct comparison of the results from the laboratory studies reviewed would be an error.

The studies used different experimental designs, different ages and visual acuity of subject populations, ambient conditions, the response variable and the manner in which it was collected were also different. The dependent variable was defined as "recognition distance" or "legibility distance." Authors often do not adequately define their response variable, making it impossible to compare across studies. The traffic engineers define legibility distance as related to either textual or symbolic signs. However, the more accepted definition of legibility distance with respect to symbol signs is *identification distance*.

It is likely that instructions given to subjects for criterion for their response were inconsistent. For example, as described above, Kline et al. (1990) invoked a "blind person criterion," whereas Bowen's study used a menu matching technique.

Zwahlen et al. (1991), in their attempts to explain why results vary so widely between field and laboratory studies, theorize the diversity in results could be a consequence of laboratory methodologies' lack of real-world fidelity. They offer several hypotheses to account for the underestimates of legibility distance obtained in laboratory studies. They believe legibility distances achieved in the laboratory are resulting from other conditions not being manipulated in the laboratory setting. The aspects of the laboratory studies under question by Zwahlen et al. are:

a. usage of monochrome monitors, projectors or visual displays to generate the target signs and visual driving scenes in driving simulators. They do not feel these devices reproduce the "real world" driving conditions with high fidelity.
b. luminances and contrasts used in laboratory studies are not specified as to how they are controlled, or what values are tested.
c. insufficient display resolution, or levels not specified.
d. small image vibrations on the monitors or projectors.
e. non-uniform and insufficiently sharp symbol or legend contours.
f. no variation in depth perception.
g. use of a dynamically changing sign and symbol while the actual viewing distance of the participant to the display stays the same.

Thus, Zwahlen et al. pose an interesting question. Which studies are reporting "reliable" legibility distances for warning symbol signs? Or, stated another way, what is the "correct" methodology to use when collecting data to be used by the Federal Highway Administration for placement guidance of traffic signs? How can the field and laboratory-based methodologies' results be reconciled?

The review above demonstrates a wide variation in recommendations which could be made based upon any one of those studies. A case can be made that field-generated data is preferable, as it has maximum face validity for application of results. Yet, researchers turn to the laboratory as a way to avoid problems inherent in the more desirable field studies. The problems with field studies include: uncontrollable variables, unpredictable events, high cost, extensive time to conduct, lack of control of illumination, trouble locating participants, and bad weather inconveniences. The laboratory affords control of variables at the expense of forfeiture of fidelity. Also, legibility distances achieved in a laboratory are suspected of being related to other conditions that are not being manipulated in the laboratory setting.

A review of the literature does not find one study which evaluates the relative effectiveness of the two methodologies, field and laboratory-based, in order to make a clear-cut choice of which to choose. Taking this assertion one step further, the present research asks the question: disregarding whether a field or laboratory study is performed, are experimental design and analyses of data without regard to within subject variability inherent in a driver having a dramatic impact on the reported results?

Special Needs of Older Drivers

The aging process affects several physical parameters involved in recognition of traffic signs. Affecting ability to perceive and read the meaning of both traffic textual and symbolic signs for the older driver are: declining static and dynamic visual acuity (Burg, 1966), longer cognitive processing times, lowering of visual processing speed (Jacewicz and Hartley, 1987; Kline and Schieber, 1982) and slower reaction time.

Advancing age also results in other physical and cognitive losses that are particularly relevant when the traffic sign is symbolic. These impairments include: degraded paced-information processing (Planek and Overend, 1973), loss of contrast sensitivity to intermediate and higher spatial frequencies (Evans and Ginsburg, 1985), longer time interval requirements for form identification (Eriksen, Hamlin, Breitmeyer, 1970), and deficits in learning capability (Schaie and Gribben, 1975). It is well established that older drivers have significantly shorter legibility distances than middle-aged and young drivers for both text and symbol traffic signs (Allen, Parseghian and Van Valkenburgh, 1984; Evans and Ginsburg, 1985; Kline, 1991; Kline, Ghali, Kline and Brown, 1990; Lum, Roberts, DiMarco and Allen, 1983; Paniati, 1988; Sivak, Olson and Pastalan, 1981).

The number of elderly persons holding a valid driver's license is increasing as the segment of persons over the age of 65 is growing. All the age-related factors described above make early detection and reading of both text and symbol traffic sign messages even more important for our growing percentage of older drivers. It is the recognition of the special needs of older drivers which may make them the design user for all highway signs. This will be discussed further as the basis for a "worst case scenario" design.

RESEARCH IMPLICATIONS

Presently, engineers who design new traffic signs or assess the placement of a current sign must use a costly and time-consuming method for evaluation of that sign. Evaluation includes *legibility distance,* as defined as the distance at which a sign can be seen and *read*, as related to words or symbol recognition (*Traffic Control Devices Handbook*, 1983, pp. 2-11). The method currently available involves either sign mock-up or a field study. However, since on-the-road studies are very costly and time- and labor-intensive, the development of laboratory measures, validated with measures taken in a driving situation, would be a major contribution to the study of traffic sign perception.

Is the decision to use a field-based versus laboratory-based methodology the only question a researcher must consider? There are other factors which are equally, if not more important to highway sign research. It appears as though there is a definite need for a standardized laboratory test which will accurately predict legibility distances in the real world for real drivers of all ages. To that end, two studies were designed and conducted where the intent was to develop such a standardized test.

Construction of a standardized test to predict legibility distances for established or new designs of traffic signs would be beneficial to the transportation community. Traffic engineers need a flexible design tool, consisting of the standardized laboratory test and a mathematical model, calibrated (validated) to adequate real-world measures, to investigate the design of new traffic signs. This model must consider the major factors affecting sign legibility distance and output optimum sign size, given constraints of conditions of use. This model, in conjunction with a standardized test administered to a limited number of people, will generate predicted field legibility distances, and thereby eliminate costly fabrication of prototype traffic signs and need for field research.

NEED FOR STANDARDIZED RESEARCH METHODOLOGY

It will be argued that recommended distances at which signs are placed to be recognized should be determined from a "worst-case" scenario. This premise requires a reexamination of our research methodologies for determining guidelines for placement of warning symbol signs. Of particular emphasis will be how many replications are needed to obtain reliable results. Replications are vital, as we know humans are enormously variable in their responses to the same stimuli - in this case distances at which drivers, especially elderly drivers, report seeing and understanding symbolic warning signs. This within subject variation, as expressed as coefficient of variation, has important implications for how researchers design experiments, methodologies employed and analyses which must be made before conclusions are drawn.

This study applied different analyses of the legibility distances collected from young and old drivers in both daytime field and laboratory studies. The research is aimed at exploring not only the experimental setting, but also implications for the methodology used to collect data.

RESEARCH OBJECTIVES OF PRESENT STUDY

The objectives of this research were the following:

a. Develop a simulation technique that addresses shortcomings found in the literature.

b. Conduct a field study and a laboratory study, using that simulation technique, of legibility distances for the same warning symbol signs with a sample of drivers widely varying in age, and equally divided between men and women. These studies, with multiple observations by each driver of each sign were to:

(1) identify important factors and their interactions affecting warning symbol sign legibility distance.

(2) provide a basis for the development of reliable evaluation techniques for new signs, suitable for implementation by traffic engineers.

PILOT STUDIES

INTRODUCTION

Before an experimental design could be finalized for the field study, one major issue was outstanding. This issue, the effect of sun angle on legibility distance, was discovered upon an examination of the research area available at the Research Annex, Riverside Campus and a literature review.

The taxi way at the Research Annex had the major benefit of possessing a uniform background that the symbol signs would be viewed against. However, because of the taxi way layout and distance required between signs, different sites were chosen along this taxi way. These different sites meant that the symbol signs would be facing in four different directions.

Bryant (1982) identifies three parameters which affect symbol sign legibility. The three are: luminance, contrast and symbol image size. A pilot study was planned and conducted on these parameters to evaluate the effects of these variables on legibility distance of the signs under investigation.

Of the three parameters, symbol image size was not a variable under control. The signs used in the study were standard-issue warning symbol signs, produced in a Texas Department of Transportation District's sign shop.

To ensure the parameters of luminance and contrast did not confound legibility distance, a pilot study was designed to evaluate the effect of sun angle, with resulting luminance differences, on sign legibility. Four sites were originally considered for the field study. All four sites were situated such that a different sun angle on the sign would result. Thus, luminance and contrast were factors which would vary, depending upon the site location for the sign. A drawing illustrating the sites and sun angles is seen in Figure 3.

Pilot Study I was designed to investigate the effects of site location and sun angle on legibility distance using one older and one younger driver. This study was performed to see if there was a sign by site interaction which would confound the analyses of sign legibility distances. One older and one younger driver were chosen so the effect of age could also be examined.

Pilot Study II was not originally planned. However, after Pilot Study I results were analyzed, a large within subject variability was found. Therefore, Pilot Study II was conducted to further examine this within subject variability. Pilot Study II used the same older subject as in Pilot Study I. A single site to view the target signs was selected to further examine within subject variability without the confounding of results attributable to a change in the luminance contrast ratio. This ratio would have changed depending upon site location chosen, in relationship to the prevailing sun angle.

Pilot Study III, using different subjects, was designed to confirm the results of the first two studies for the main effect of site location, a sign by site interaction, as well as the within subject variability on legibility distance.

It must be noted that sun angle is only relevant to the effect it would have on the luminance and contrast of symbol signs themselves. The pilot studies never occurred early enough in the morning nor late enough in the early evening to have sun angle be a problem for

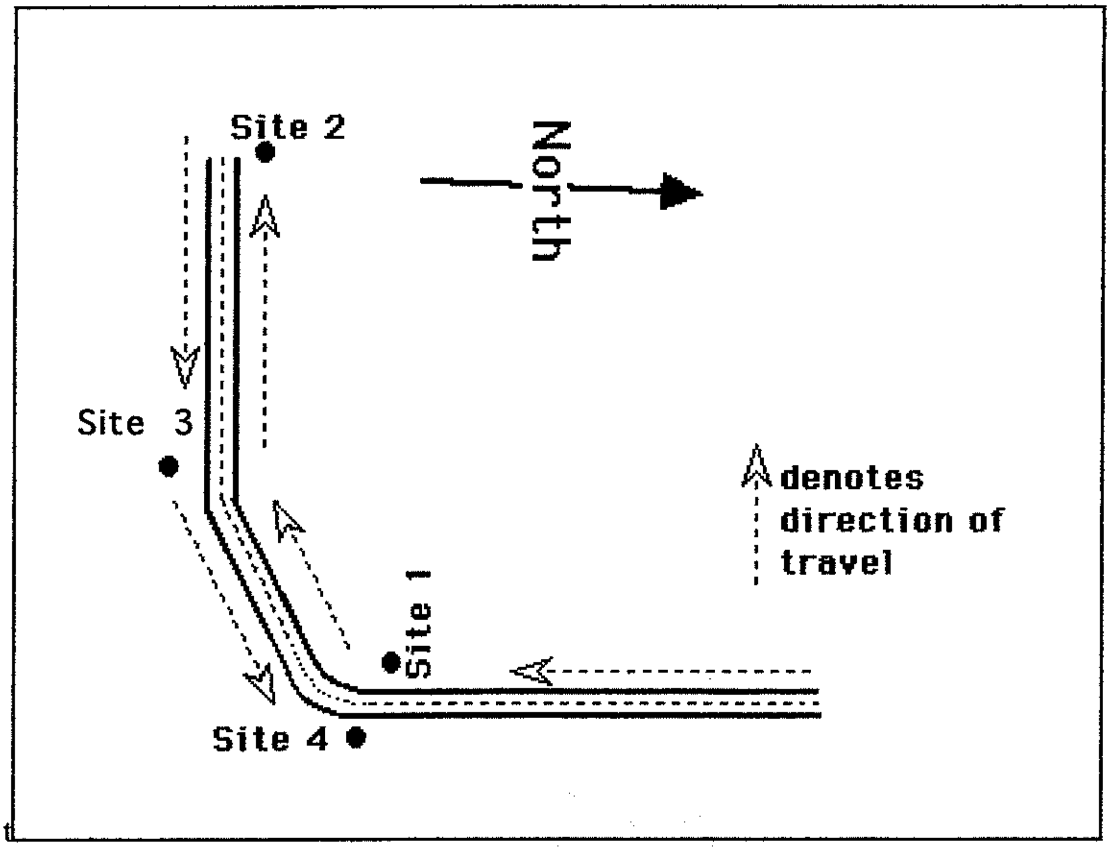

Figure 3. *Schematic of four site locations and sun angle, Riverside Campus*

he drivers' eyes. These studies were all conducted during the months of May and June between the hours of 8:30 a.m. and 4:00 p.m. The sun was never low enough in the sky at any of those times to be an obstacle to the drivers' viewing the signs. At no time was the sun directly shining into the eyes of the driver.

The consideration for sun angle was related to the effect of the sun reflecting off the retroreflective material surface of the warning sign. All road signs are constructed from a special material which reflects headlamp light to aid in their detection, recognition and legibility at night. This reflection from the sun during the day would change one of Bryant's (1982) factors affecting legibility distance, that of luminance contrast.

PILOT STUDY I

METHODOLOGY

Pilot Study I was performed to assess the effect of sun angle and resulting contrast ratio of symbol and background on legibility distance of the signs mounted on four site locations as illustrated in Figure 3 above on.

Two subjects were chosen, an older male, aged 64 with a visual acuity of 6/6 (20/20) and a young female, aged 24, with a visual acuity of 6/7.4 (20/25). Two signs were used in Pilot Study I: *Bicycle Crossing* (W11-1) and *Crossroad* (W2-1). These two signs represent the extremes of complexity of the six signs investigated by Bowen (1993). The *Crossroad* symbol is a simple design, composed of thick, bold lines. *Bicycle Crossing* has complex detail and uses thin lines to convey the detail. Figures 4 and 5 illustrate the warning symbol signs used in Pilot Study I.

Figure 4. *Bicycle Crossing (W11-1) warning symbol sign*

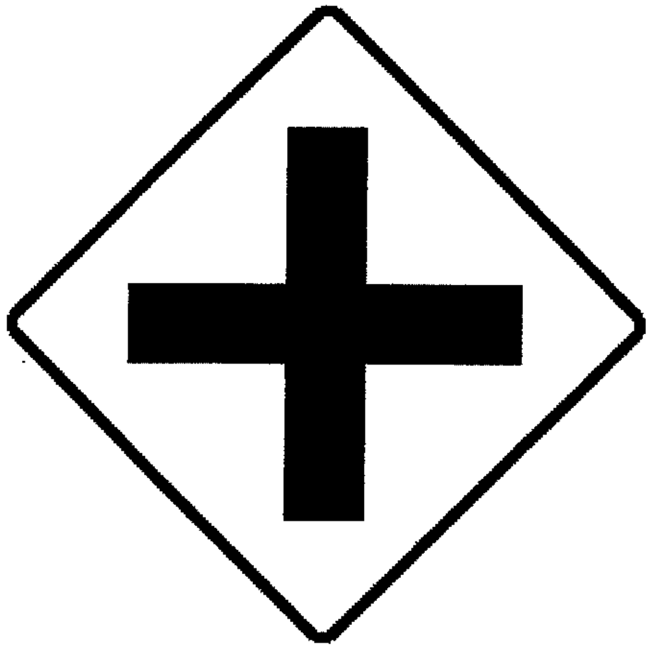

Figure 5. *Crossroad (W2-1) warning symbol sign*

To prevent sign conspicuity from confounding legibility distance, a uniform background against which signs were viewed was needed. A background of trees and grass was available along a stretch of taxi way at the Riverside Campus. Along this taxi way where the grass and treed background was available in all directions, the four sites were chosen, which took advantage of the signs always appearing against the trees.

From Figure 3, it can be seen that signs at site 1 face a northerly direction, and thus, the sun does not directly illuminate the sign face. Site 2 signs face easterly, where the sun would shine directly on the face in the early morning hours. Site 3 has the signs mounted facing a westerly direction, so the late afternoon/early evening sun would shine on the sign face. Site 4 signs are facing south, where like site 1, the sun never directly illuminates the face of the sign.

All locations have the same background against which the warning symbol signs were seen - trees. The background was standardized so as not to affect conspicuity of the sign against different backgrounds and influence legibility distances.

Signs were mounted on a wooden quick-change pole device, specifically constructed for this research. Figures 6 and 7 show the mounting pole and its operation. The pole placed the sign at the MUTCD recommended height of 2.1 meters (7 feet) from the pavement to the bottom of the sign. The lateral displacement of the sign was 1.83 meters (6 feet) from the edge of the traveled way.

Figure 6. *Quick-change mounting pole for signs*

Figure 7. *Operation of quick-change mounting pole for signs*

Procedure

Ten replications were made of each sign using all four locations. Nine additional signs were used for distractors to discourage guessing of the two signs of interest. The nine distractor signs were:

a. *Winding Road Right* (W1-5R)
b. *Turn Sign Right* (W1-1R)
c. *Turn Sign Left* (W1-1L)
d. *Reverse Turn Right* (W1-1R)
e. *Y-Intersection* (W2-5)
f. *Side Road 45 Degrees* (W2-3)
g. *Side Road 90 Degrees* (W2-2)
h. *T-Intersection* (W2-4)
i. *Deer Crossing* (W11-3)

These nine distractor signs were chosen from the remaining 17 signs acquired from the Texas Department of Transportation. *Deer Crossing* was the only sign which had thin lines and served as the best distractor for *Bicycle Crossing*. The other eight signs were meant to be confused with *Crossroad*, while others simply served as diversions from the two signs of main interest.

A random order of presentation of the 11 signs was produced. A total of 56 trials were collected, consisting of ten replications of *Bicycle Crossing* and *Crossroad* and four each replications of the nine distractor signs. No more than four replications were collected on the distractor signs. The logistics of moving signs for additional exposures would have slowed down the experimental session such that completion would have been excessive, leading to fatigue on the part of the participants.

Pre-training on 27 potential warning signs was given. A menu containing 27 warning symbol signs was reviewed and training accomplished until each subject achieved a 100% correct criterion when asked to identify the signs presented on the menu. Instructions given the subject were that he/she was to respond by number on the sign menu when identity was "certain." Certainty was defined as expressed willingness to make a vehicle maneuver in response to the warning sign as perceived. The sign menu was available to the participants at all times during the experimental session. The task was one of matching their perception of the sign in the field with the symbol sign on the menu. Subjects were informed that the signs they would encounter would always be one of the 27 signs on the menu. However, they were cautioned that they might not see all 27 signs and some signs might be seen more than once. Subjects responded to all signs, regardless of the signs being distractor or the two of interest.

Subjects drove a Honda Accord at 15 miles per hour to facilitate the assistant's job of changing warning signs at all four locations. When the subject was certain as to the sign's identity, a bag with a trial number designation was dropped out the window of the car. All sign trials were treated exactly the same. In other words, a bag was dropped for all trials. The distractor sign trials had bags with a different color sand/gravel for ease of identification. The trials on the two signs of interest were color coded, as well as containing a trial number from one through 10.

After all trials were collected, a measuring wheel was used to compute the distance from the sign to the point at which it was identified. This distance was sign legibility, the dependent variable.

RESULTS

A 2 X 2 X 4 analysis of variance (ANOVA) was performed on the mean legibility distance in order to assess the effects of subject, sign, and site location. Table 5 is the source table from that ANOVA.

TABLE 5

ANOVA source table for Pilot Study I

Source	df	Sum of Squares	F Ratio	Prob >F
Sign	1	1417857.2	81.26	0.0000*
Subject	1	20049.6	1.15	0.2940
Sign*Subject	1	113795.7	6.52	0.0170*
Site	3	79053.9	1.51	0.2370
Sign*Site	3	118679.1	2.27	0.1070
Subject*Site	3	13549.1	0.26	0.8540
Sign*Subject*Site	3	60476.8	1.16	0.3470

*** denotes significance at alpha = 0.05**

Only the two signs of concern, *Bicycle Crossing* and *Crossroad* were evaluated. There was a significant main effect for sign and a significant interaction between sign and age group. As might be expected because of differences in static visual acuity, the younger subject had a longer mean legibility distance than the older driver, although there was no significant main effect of subject (young = 246.6 meters or 809 feet; older = 234 meters or 767.7 feet). For sign, *Bicycle Crossing* had a mean legibility distance of 181.9 meters (596.9 feet), while *Crossroad* had a mean legibility distance of 296.6 meters (973.1 feet).

The two-way interaction between sign and age is shown graphically in Figure 8. The mean legibility distance for the younger driver for *Bicycle Crossing* was 173.1 meters (567.9 feet), while the older driver saw it further away - 190.6 meters (625.4 feet). The opposite pattern is seen for legibility distance of the *Crossroad* sign: the older driver's mean distance was 277.4 meters (910.0 feet); while the younger driver had a longer legibility distance of 320.1 meters (1050.1 feet).

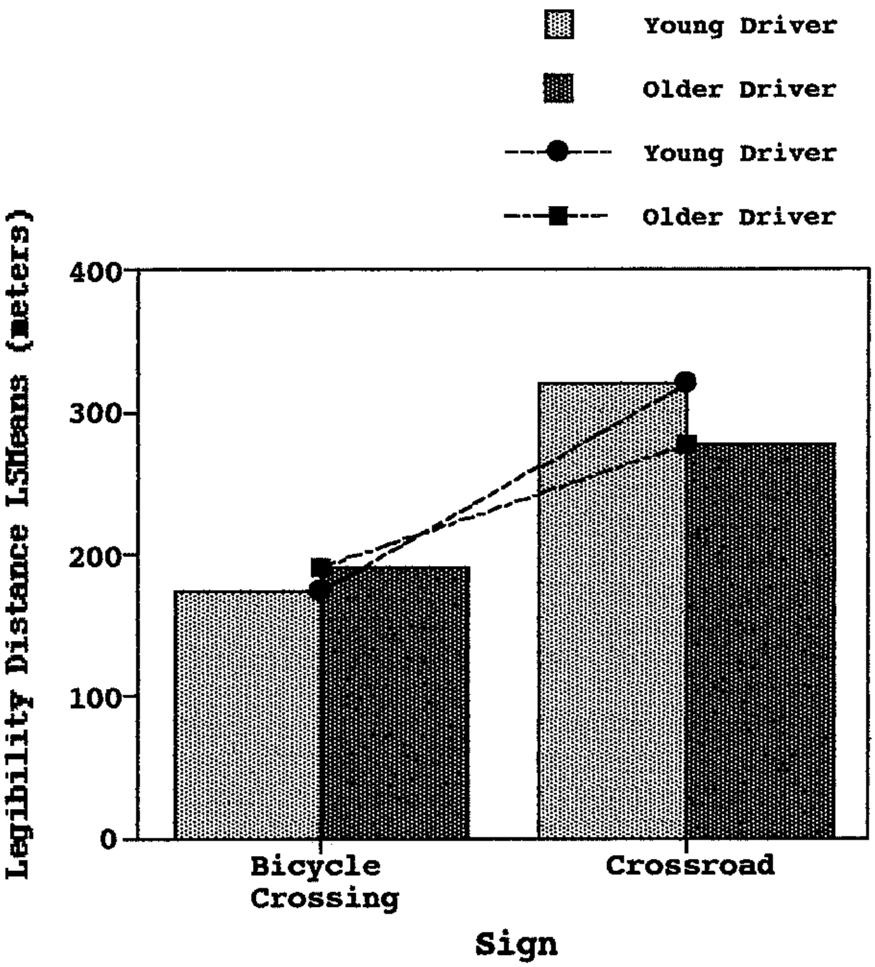

Figure 8. *Plot of Sign*Age interaction for Bicycle Crossing and Crossroad sign*

The main effect of site location was not significant. This meant that for the four locations chosen, the sun angle on the sign, resulting in a change in luminance and contrast ratio, had no effect on legibility distance for either *Bicycle Crossing* or *Crossroad* signs. Remember that the hours of data collection never resulted in the sun angle being directly in the eyes of the driver. The sun was never any lower than 30-40° from the horizon. For the four sites, the mean legibility distances were the following:

Site 1	248.2 meters (814.4 feet)
Site 2	251.8 meters (826.0 feet)
Site 3	228.5 meters (749.6 feet)
Site 4	228.4 meters (749.3 feet)

There was also no Sign X Site interaction. So for the two extremes of complexity of symbol signs investigated, the site location had no effect on legibility distance. By site number, the following are the legibility distances for *Bicycle Crossing:*

Site 1	202.6 meters (664.6 feet)
Site 2	190.5 meters (624.9 feet)
Site 3	162.8 meters (534.2 feet)
Site 4	171.7 meters (563.2 feet)

Examining the same mean legibility distances by site location for *Crossroad*, the results are:

Site 1	287.6 meters (943.7 feet)
Site 2	319.1 meters (1047.1 feet)
Site 3	291.4 meters (955.9 feet)
Site 4	288.2 meters (945.5 feet)

A major finding from Pilot Study I was the large within subject variability for legibility distance. For the young subject, the coefficient of variation (CoV) for legibility distance, defined as:

$$[(\text{standard deviation} \div \text{mean legibility distance}) * 100]$$

expressed as a percentage, for *Bicycle Crossing* was 24.2% (151.7/625.4), and was 12.9% for *Crossroad* (117.1/910.0).. The older driver had very similar within subject variability. For *Bicycle Crossing* the CoV was 23.2% (131.4/567.9). For the same older driver, the CoV for *Crossroad* was 14.3% (150.2/1050.1). In the case of both ages of drivers, the within subject variability, expressed as CoV, is almost twice as large for the *Bicycle Crossing* sign as for *Crossroad.* However, for both subjects, the CoV for both signs is large. The CoV is based on 10 observations taken for each of the two signs.

The sign with the most complex small detail, *Bicycle Crossing*, had a shorter legibility distance for the older driver. This is not too surprising, since there is a loss in sensitivity to higher spatial frequency detail with increasing age (Ginsburg, Evans, Cannon, Owsley and Mulvanny, 1984, and Owsley, Sekuler and Siemsen, 1983). However, the same older driver saw the simple, bold *Crossroad* sign at a further distance than the young subject. This could be explained by the fact that the corrected visual acuity for the older driver was 6/6 (20/20), but was 6/7.5 (20/25) for the younger driver.

CONCLUSIONS AND IMPLICATIONS FOR PILOT STUDY II

Since the main effect of site location was not significant, only one site was chosen for Pilot Study II and for the field study portion of this research. The times of experimental data collection were such that the sun was never in a position in the sky where it was shining directly into the drivers' eyes during any data runs. It was also known at this time that the field data collection would occur only between the hours of 8:00 a.m. at the earliest, and 4:30 p.m. at the latest.

Referring to Figure 3, site 1 was the position used for the second pilot study and also for the field study. Site 1 always had the sun behind the sign throughout the day. In other words, the sun never illuminated the face of the symbol sign.

PILOT STUDY II

METHODOLOGY

Based upon the results of the first pilot study where large within subject variability was found, a second pilot study was executed using the same older driver from Pilot Study I. The purpose of Pilot Study II was to further investigate the existence of large within subject variability evidenced in the first pilot study. Two different target signs were chosen from those of primary interest in Pilot Study I. *Divided Road Ends* and *Narrow Bridge*, represented the extremes of complexity based upon Bowen's (1993) analysis of spatial frequency content. *Narrow Bridge* had high complex spatial frequency content; *Divided Road Ends* was a less complex sign with low spatial frequency content. This choice of signs was meant to parallel the complexity and spatial frequency content of signs investigated in Pilot Study I - *Bicycle Crossing* and *Crossroad.* Figures 9 and 10 show the two signs of interest.

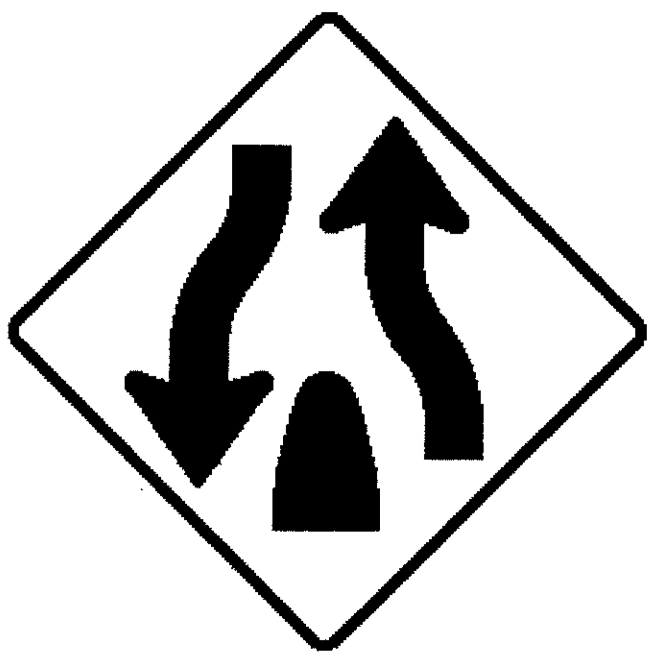

Figure 9. *Divided Road Ends (W6-2) warning symbol sign*

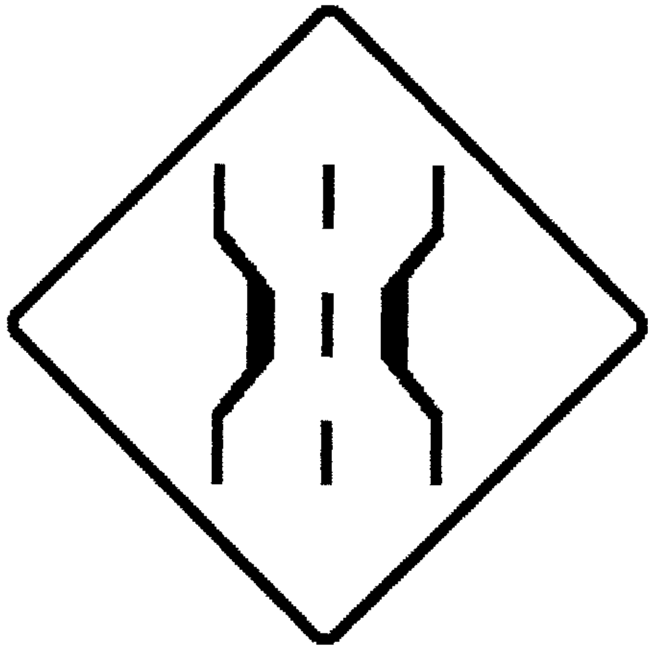

Figure 10. *Narrow Bridge (W5-2a) warning symbol sign*

The purpose of Pilot Study II was twofold: first to examine whether the large within subject variability for the two signs investigated in the first pilot study could be attributed in part to site location. Therefore, only site 1 was used, where the sign is facing northward so the sun never illuminated the sign face. This site was selected to eliminate any interaction with contrast ratio as a function of shadows and sun location relative to the sign. A discussion of this location and findings from Pilot Study I are found above. Pilot Study II was conducted between the hours of 9:30 a.m. and 12:00 p.m.

As noted above in the first pilot study, a total of ten replications per sign per subject were collected. However, those ten replications per sign were spread out over four sites. Therefore, a subject saw the same sign at each of the four sites only two, or at most three times (since 10 replications divided over the four locations means each sign could be seen at most 2.5 times - hence, the two or three exposures to each sign at each location). The large within subject variability discovered in Pilot Study I may be due to a confounding of sign with site location, where the sun hits the sign from different directions. The small number of observations of each sign at each location was insufficient to draw any conclusions about within subject variability.

The second objective of Pilot Study II was to collect a larger number of trials (10) from one subject at a standardized location. The older driver was selected due to his familiarity with the task and because older people tend to have more within subject variation than younger.

Procedure

Ten days had elapsed since Pilot Study I, therefore, a training session was conducted again with the older driver. This was meant to familiarize him with the menu of symbol signs.

A total of 50 trials were collected, 10 replications on each of the two target signs, plus three replications each on 10 distractor signs. The ten distractor signs were:

a. *Bicycle Crossing* (W11-1)
b. *Deer Crossing* (W11-3)
c. *Hill Sign* (W7-1)
d. *Curve Sign Right* (W1-2R)
e. *Right Lane Ends, Merge Left* (W4-2)
f. *Reverse Turn Right* (W1-4R)
g. *Slippery When Wet* (W8-5)
h. *Y-Intersection* (W2-5)
i. *T-Intersection* (W2-4)
j. *Side Road 45* (W2-3)

The identical procedure outlined for Pilot Study I was used in Pilot Study II. The same menu of 27 signs was available at all times for the subject to match his perception. The subject drove at the same speed approaching the sign and used the same response criterion, certain response, as in the first pilot study.

RESULTS

Only the data for the two signs of interest, *Divided Road Ends* and *Narrow Bridge* were analyzed. A coefficient of variation (CoV) was computed in the same fashion as described above for Pilot Study I. For *Divided Road Ends*, the CoV was 25.75% (225.1/874.8). *Narrow Bridge* had a CoV of 25.96%.

As expected, the *Divided Road Ends* sign was recognized from a further distance - 266.6 meters (874.8 feet). *Narrow Bridge* had a mean legibility distance of 233.6 meters (766.5 feet). A t-test on the means reveals that they are not statistically significant ($t(18) = 1.141$, $p < 0.269$). It is possible that using Bowen's (1993) Legibility Figure of Merit for spatial frequency content did not distinguish sufficiently between the two signs. A complex and non-complex sign were required to be tested. It is likely that the sign *Divided Road Ends* was not as "non-complex" as *Crossroad*. Therefore, the magnitude of differences found in Pilot Study I between the two signs were not found here. The two signs of interest may not have, in actuality, been that different in complexity. *Narrow Bridge* is composed of long, thin parallel lines, which may have made it more distinguishable (and hence more legible) by its lack of definitive shape.

CONCLUSIONS

Narrowing the site location to only one meant the signs viewed were northerly facing, with the sun never shining directly on the sign face, no matter what time of day. This viewing location did not reduce the within subject variability of legibility distances for the two different signs tested. The within subject variability was a robust finding, and bore further investigation.

The results showed that for site 1 where the sun angle was such that the sun never directly illuminated the sign, large within subject variability was still present. The within subject variation was even larger than that found in Pilot Study I for this older subject. Even *Divided Road Ends*, which was a less complex sign than *Narrow Bridge*, had a large CoV, being only 0.21% less than for *Narrow Bridge*. At this point, no speculations were made as to what caused the within subject variability.

Ten replications of the two signs appeared to be sufficient to validate the within subject variation evident for both subjects in Pilot Study I.

PILOT STUDY III

METHODOLOGY

The main effect of site location was investigated in a third pilot study. If there was no effect of site location on legibility distance for warning symbol signs chosen, then one site could be used for field data collection. Excess time to complete a set of trials was noted during Pilot Study I. The logistics of transferring signs to four different sites lengthened the session tremendously. The older driver noted his fatigue during the session. In addition, it was felt that the sun angle would change excessively during the time needed to move single copies of signs to four different sites. For all these reasons, one site to view all signs would be preferable for the field study.

Three new, young subjects (two males and one female) aged 22, 23 and 27, served as participants. Two sites were chosen for sign placement - sites 1 and 2, as seen in Figure 3. These two sites ensured there was no straight-on illumination of the sign face in one case (site 1), and the sun was reflecting off the front of the sign in the mid-morning for the second. All trials were collected during mid-July 1993, before 12:00 p.m. to control for sun angle on site 2. It should be noted that with daylight savings time in effect, the sun angle was comparable to 11:00 a.m. At no time did any subject drive toward site 2 with the sun shining directly in his/her eyes.

Subjects' visual acuity was screened using a Snellen chart viewed from 6.1 meters (20 feet). Contrast sensitivity was tested using the Vistech Consultants Inc. Vision Contrast Test System, VCTS (Vistech Consultants, 1988) viewed from 10 feet. The resulting visual acuity and contrast sensitivity measures for the three pilot subjects are presented in Table 6.

TABLE 6

Visual acuity and contrast sensitivity measurements for Pilot Study III subjects

Subject (Age)	Visual Acuity Left Eye in meters (feet)	Visual Acuity Right Eye in meters (feet)	Contrast Sensitivity Equivalent Acuity Values
Subject Male (22)	6/3.9 (20/13)	6/6 (20/20)	6/4.5 (20/15)
Subject Female (23)	6/9 (20/30)	6/9 (20/30)	6/9 (20/30)
Subject Male (27)	6/7.5 (20/25)	6/9 (20/30)	6/4.5 (20/15)

Signs of principal interest were *Bicycle Crossing* (W11-1) and *Crossroad* (W2-1), as illustrated above in Figures 4 and 5. These signs were chosen for the reasons stated in Pilot Study I. Ten replications were collected on each of the two signs, five at each of the two site locations per subject. A total of 46 trials were collected; the remaining 26 were for seven distractor signs. The seven distractor signs were:

a. *Two-Way Traffic* (W6-3)
b. *Divided Road Ends* (W6-2)
c. *Hill Sign* (W7-1)
d. *Narrow Bridge* (W5-2a)
e. *Slippery When Wet* (W8-5)
f. *Y-Intersection* (W2-5)
g. *T-Intersection* (W2-4)

The same pre-training and procedure outlined for Pilot Study I were used in this pilot study.

RESULTS

Within subject variation was examined on individual subjects for the two signs. Table 7 summarizes the CoV for both signs for the three subjects.

TABLE 7

Summary of Coefficient of Variation for *Crossroad* and *Bicycle Crossing* for three subjects in Pilot Study III

Subject	**Coefficient of Variation for *Crossroad***	**Coefficient of Variation for *Bicycle Crossing***
#1 - Age 22 (Male)	8.1% (130.2/1595.4)	26.7% (334.5/1254.6)
#2 - Age 23 (Female)	15.8% (136.7/863.9)	32.4% (143/441/9)
#3 - Age 27 (Male)	17.5% (222.9/1275.5)	34.8% (375.6/1078.4)

For all subjects, the CoV for *Bicycle Crossing* is much larger than that for the *Crossroad* sign. CoV for *Bicycle Crossing* was nearly two times as large as the CoV for *Crossroad* for subject 3. The CoV for *Bicycle Crossing* was over twice as large as the CoV for *Crossroad* for subject 2. Subject 1's CoV for *Bicycle Crossing* was over three times larger than for *Crossroad*. Note that subject 1 had the smallest CoV for *Crossroad* at only 8.1%. This subject had the best visual acuity of 6/3.9 (20/13) of the three subjects (See Table 7). Subject 1 also had a contrast sensitivity visual acuity equivalent of 6/4.5 (20/15). The other two subjects had less than 6/6 (20/20) static visual acuity in both eyes.

In order to verify the effect of site location on legibility distance a 4 X 2 ANOVA was performed on the mean legibility distances to ascertain the effects of site and sign. Table 8 shows the source table for the ANOVA test.

TABLE 8

ANOVA source table for Pilot Study III

Source	df	Sum of Squares	F Ratio	Prob >F
Subject	2	2562128.7	20.8092	0.0000*
Site	1	58257.0	0.9463	0.3355
Subject*Site	2	55912.9	0.4541	0.6377
Sign	1	654677.2	10.6344	0.0020*
Subject*Sign	2	117760.4	0.9564	0.3915
Site*Sign	1	145991.2	2.3714	0.1301
Subject*Site*Sign	2	54958.5	0.4464	0.6426

*** denotes significance at alpha = 0.05**

CONCLUSIONS AND IMPLICATIONS FOR FIELD STUDY

Even using young drivers as subjects, Pilot Study III showed large amounts of within subject variability. The same signs were investigated in both Pilot Studies I and III. The CoV for *Crossroad* was comparable between the two studies, ranging between 8.1% and 17.5%. *Bicycle Crossing*'s CoV ranged between 23.2% and 34.8%.

From Table 8 above, it can be seen that only the subject and sign variables were significant. Figure 11 and Figure 12 show the plots for the main effect of subject and sign respectively.

The mean legibility distance collapsed over both signs for subject 1 is 434.3 meters (1425 feet), 199.0 meters (652.9 feet) for subject 2, and 358.7 meters (1177 feet) for subject 3.

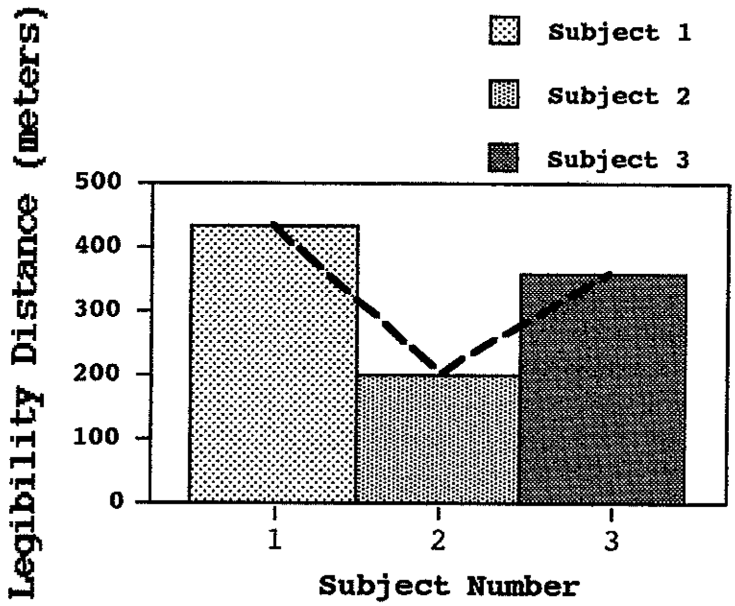

Figure 11. *Main effect plot of subject for Pilot Study III*

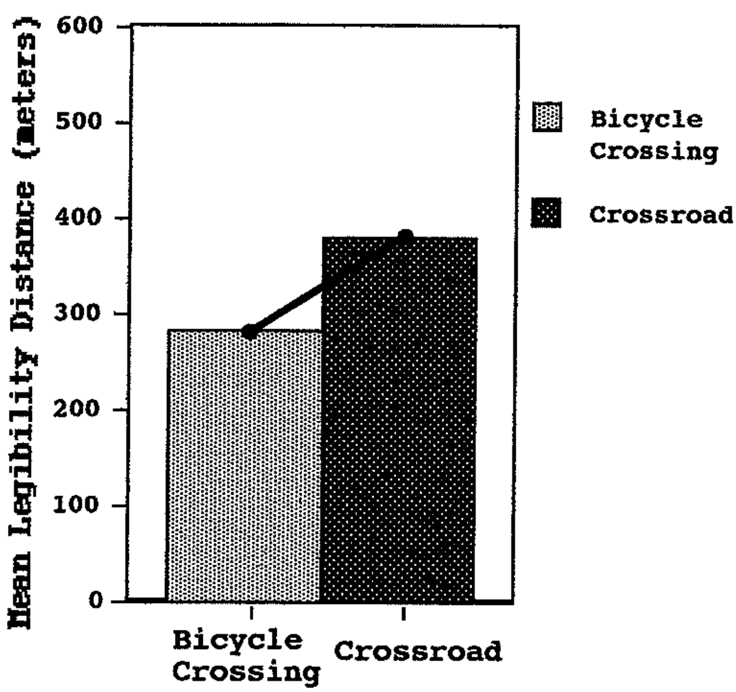

Figure 12. *Main effect plot of sign for Pilot Study III*

For the sign *Bicycle Crossing*, the mean legibility distance for all subjects was 281.9 meters (925 feet). *Crossroad* had a mean legibility distance of 379.5 meters (1245 feet). *Crossroad* was seen 1.35 times farther than *Bicycle Crossing* for these subjects in this pilot study.

The Site X Sign interaction was not significant ($p < 0.1301$). Neither the legibility distance for a complex sign, *Bicycle Crossing*, nor a simple, bold sign, *Crossroad*, was influenced by the site location. Recalling that sites 1 and 2 differed in that site 2 had the sun on the sign face during the entire test. This location was affected maximally by sun angle and ensuing luminance and contrast ratio changes. Site 2 represented the worst case for sign viewing, yet it did not affect legibility distance for signs viewed at that location, versus the same signs seen at site 1.

The main effect of site not being a significant independent variable, coupled with the non-interaction of site location with sign legibility distance, had two implications for the field study. First, the site location was not contributing directly to the large within subject variability. By testing the most and least complex warning symbol signs (Bowen, 1993), the non-significant Sign X Site interaction indicated site location did not affect legibility distance for either sign.

Therefore, it made no difference which site was chosen for the field study. Site 1 was selected because the sun never directly illuminated the sign face. Other logistics considerations were also optimized by this site selection. The terrain surface was flat for the entire distance along the taxi way approaching site 1. Sites 2 and 3 had rises and dips in the terrain. Additionally, the issue of excessive subject fatigue and change of sun angle were also avoided with the single site.

The within subject variability became a topic of major interest for the field study evaluation. A sensitivity analysis was performed on the 10 trials of legibility distance collected for both signs in Pilot Study III. It was determined from this analysis that although the within subject variability for legibility distance was large, it became stable after the fifth trial, with the standard deviations becoming smaller. Based upon the results of this sensitivity analysis, six replications per sign were chosen for the field study.

SUMMARY OF EXPERIMENTS, DESIGN AND ANALYSES FOR FIELD AND LABORATORY STUDIES

EXPERIMENT

Statement of Problem

Field Study: Using a wide range in age of subject population, establish a baseline performance for the response variable, legibility distance, for six target warning symbol signs. In addition, identify important factors and interactions that affect sign legibility distance.

Laboratory Study: Develop a new simulation technique based upon the criticisms of Zwahlen et al. (1991), to collect legibility distances for the same warning symbol signs.

Choice of Response or Dependent Variable

The dependent variable for both field and laboratory studies is legibility distance. This response is defined as the distance at which a participant is certain as to the identity of the symbol sign. The participant responds by matching his/her perception of the sign from a menu of 27 possible signs. The participant answers by giving the number from the menu which matches the sign they see. They must be certain as to the sign identity, such that they are willing to make a vehicle maneuver in response to the sign message.

In the field study, the dependent variable is measured in feet by a fifth wheel measuring device. In the laboratory study, the distance is collected in a computer program which keeps track of the frame number in the animation where the sign match is made. The frame number equates to a distance the participant is from the sign.

Selection of Factors and Levels To Be Varied

The independent variables are gender, age category of subject and individual sign. The factors of gender and age category are nested within subject. Gender is obvious; age is defined as older and young drivers. By "older", the subject must be ≥ 65. "Young" drivers must be under the age of 40. This is to control for presbyopia, a very common occurrence at age 40. Substantial changes in visual acuity and contrast sensitivity are seen at and above age 65.

Twelve subjects are in each group, for a total of 24 subjects. Gender is equally divided within the two age groups, to the extent possible.

Age and gender are between subject variables and are random. Subject[Gender,Age] is, of course, a qualitative variable, being nominal. Sign is a fixed factor. A total of six warning symbol signs are of primary concern as stimuli (independent variables). The six signs are chosen using candidate signs' legibility figure of merit (LFOM), a concept developed and tested by Bowen (1993). LFOM is related to the spatial frequency content of the symbol sign and is indexed against age and contrast sensitivity curves (Owsley, Sekuler and Siemsen, 1983). Two signs are selected, each representing low, medium and high LFOM. Condition is a fixed factor with two levels: field and laboratory. These are discussed below.

How Factor Levels Are Combined

The field and laboratory studies are repeated measures designs. All subjects serve in both conditions. All subjects observe the same six signs.

DESIGN

Number of Observations To Be Taken

All 24 subjects forming the two age groups observe the six target signs six times for both the field and laboratory study. An additional number of signs (12) are used as distractors in both field and laboratory studies so the subjects will not know which of the sign trials are of primary interest. This is also to control for learning. In both studies 24 additional trials on the 12 signs will be collected, for a total of 60 trials per subject. However, only the six target sign data will be part of the analyses.

Order of Experimentation and Method of Randomization

Due to logistics considerations, all subjects run in the field condition of the research first. The field study is to establish baseline legibility distances. This part of the experiment must be conducted first for all subjects.

Using a table of random numbers, a sequence is constructed for the 60 trials. All subjects receive the same random order of signs. Since the twelve signs used in the field study as distractors are not the same as for the laboratory condition, a different sequence of 60 random signs is generated. Again, all subjects see the same sequence in the laboratory study.

Mathematical Model to Describe the Experiment

The following is the model used including condition, field or laboratory, as an independent variable:

$$\hat{Y}_{i(jk)lmo} = \mu + \alpha_j + \beta_k + \theta_l + \Omega_m + (\alpha\beta)_{jk} + \delta_{i(jk)} + (\alpha\theta)_{jl} + (\alpha\Omega)_{jm} + (\beta\theta)_{kl} + (\theta\Omega)_{lm} + ((\beta\Omega)_{km} + (\alpha\beta\theta)_{jkl} + (\alpha\beta\Omega)_{jkm} + (\alpha\theta\Omega)_{klm} + (\beta\theta\Omega)_{klm} + (\alpha\beta\theta\Omega)_{jklm} + (\theta\delta)_{i(jk)l} + (\Omega\delta)_{i(jk)m} + (\theta\Omega\delta)_{i(jk)lm} + (\theta\Omega\delta)_{i(jk)lm} + \varepsilon_{i(jk)lmno}$$

Where: μ = population mean

α_j = effect of the j-th level of age category
j = 1-Young (<30), 2-Older ($\geq$65);

β_k = effect of the k-th level of gender -
k = 1-male, 2-female;

θ_l = effect of the l-th level of treatment-
l = 1-field condition, 2-laboratory condition

Ω_m = effect of the m-th level of treatment -
m= 1,2,3,4,5,6 - individual sign

$(\alpha\beta)_{jk}$ = interaction between j-th level of age category and the k-th level of gender

$\delta_{i(jk)}$ = effect of subject nested within the j-th level of age and the k-th level of gender

$(\alpha\theta)_{jl}$ = interaction between j-th level of age category and the l-th level of condition

$(\alpha\Omega)_{jm}$ = interaction between j-th level of age category and m-th individual sign

$(\beta\theta)_{kl}$ = interaction between k-th level of gender and the l-th level of condition

$(\beta\Omega)_{km}$ = interaction between k-th level of gender and the m-th individual sign

$(\theta\Omega)_{lm}$ = interaction between l-th level of condition and the m-th individual sign

$(\alpha\beta\theta)_{jkl}$ = interaction between j-th level of age category and the k-th level of gender and l-th level of condition

$(\alpha\beta\Omega)_{jkm}$ = interaction between j-th level of age category and the k-th level of gender and m-th individual sign

$(\alpha\theta\Omega)_{klm}$ = interaction between k-th level of age category, the l-th level of condition and the m-th individual sign

$(\beta\theta\Omega)_{klm}$ = interaction between k-th level of gender, the l-th level of condition and the m-th individual sign

$(\alpha\beta\theta\Omega)_{jklm}$ = interaction between the j-th level of age category and the k-th level of gender and the l-th level of condition and the m-th individual sign

$(\theta\delta)_{i(jk)l}$ = interaction of l-th level of condition with subjects nested within age category and gender

$(\Omega\delta)_{i(jk)m}$ = interaction of m-th individual sign with subjects nested within age category and gender

$(\theta\Omega\delta)_{i(jk)lm}$ = interaction of l-th level of condition with m-th individual sign with subjects nested within age category and gender

$(\theta\Omega\delta)_{i(jk)lm}$ = interaction of l-th level of condition with m-th individual sign, with subjects nested within age category and gender

$\varepsilon_{i(jk)lmo}$ = sampling error, o = 1-6 replications

Hypotheses To Be Tested

The null hypothesis under test is that there is no difference between subjects' performance on legibility distance to the six signs between conditions, field and laboratory. Additional null hypotheses are that there are no effects due to any of the factors varied: age, gender, sign or condition or any interactions.

ANALYSIS

Data Collection and Processing

In the case of the field study, the data collection is performed using a fifth wheel measuring device, a digital distance meter readout and a video camera. The video camera tapes the digital readout. The audio channel on the tape is used to signal the experimenter as to when the subject's response occurs. The distance is then read from the readout box on the tape.

For the laboratory study, all data collection is automated through the Macintosh computer, whereby the experimenter inputs a key stroke corresponding to the subjects' response. If it is incorrect, the trial continues. If it is correct, the distance is recorded in a file for that subject and that sign observed.

Computation of Test Statistics

Due to the fact that this design has both fixed and random factors, an Analysis of Variance (ANOVA) will be run using the procedure PROC MIXED to analyze the data (SAS Institute, Cary, NC.). According to Hicks (1982), the repeated measures design is but a special case of the nested-factorial design. The study has repeated measures in that all subjects serve as their own control, i.e., the same subjects are used in both conditions - field and laboratory and view all six target signs.

The *between* subject factors are age category (young or older), gender (male or female), the interaction of age with gender, and subject nested within age and gender.

The *within* subject factors are the main effects of condition (field or laboratory) and sign (*Crossroad*, *T-Junction*, *Slippery When Wet*, *Two-Way Traffic*, *Bicycle Crossing*, and *Deer Crossing*). These main effects are crossed with the between subject factors, as seen in the mathematical model above.

In addition, Spearman's Rank Correlation will be used to see if the ranks of signs by legibility distance differ between age, gender, and condition.

Additional post hoc statistics will be applied, if applicable.

Interpretation of Results

The results will be interpreted in light of the past research studies which examined the same symbol signs' legibility distances. The new laboratory simulation technique will be evaluated to see if it can provide a basis for the development of a refined, reliable evaluation technique for new signs, suitable for implementation by traffic engineers.

FIELD STUDY

INTRODUCTION

One of the objectives of this research, as discussed in the section Introduction, was to conduct a field and laboratory study (using a newly developed simulation technique) of legibility distances for the same warning symbol signs with a sample of drivers widely varying in age, and equally divided between men and women. The purposes of this field study were to:

(a) identify important factors and their interactions affecting warning symbol sign legibility distance.
(b) provide a basis for the development of reliable evaluation techniques for new signs, suitable for implementation by traffic engineers.

This field study was the first phase of the research needed to accomplish (a) and (b) above.

METHODOLOGY

The results and lessons learned from the three pilot studies outlined above helped in refining procedure, methodology, and experimental design for the field study. Out of the three pilot studies the following was learned:

a. with the times of day during when data collection was to occur, site location was not a contributing factor to legibility distance. Since the sun would never be shining in the eyes of the driver and the sign face would not be illuminated directly by the sun, a single site, site 1, was chosen.

b. controlling for factors like sign complexity and site location, within subject variability was a very apparent phenomenon. The mention of within subject variability had never before been encountered in the literature review accomplished for this research. The lack of authors' usage of replications of response variable is very likely to be the reason why previous researchers never discovered this variability. It is well known that humans can be variable in their responses to the same stimuli. But the magnitude of the variation cannot be calculated and evaluated without the use of replications. Replications were mandatory in this study to investigate subjects' variability.

Research Participants

The same twenty four drivers participated in both the field and laboratory studies. None had participated previously in any of the pilot studies. The 24 participants were divided into two equal-sized age groups: older and Young. The older group consisted of an equal number of men and women (six each). The young group had seven males and five females. Table 9 shows the mean age for both groups, as well as the age range. All participants held current Texas drivers' licenses and were active drivers. Each subject drove in this study with corrective lenses, if they normally drove a car with vision correction.

The older group members were paid $30 for their time to perform both field and laboratory studies, but only after completion of the laboratory portion. The young group consisted of graduate and undergraduate students who received course credit for their participation.

Visual Screening of Participants

A Snellen chart viewed from 6.1 meters (20 feet) was used to measure corrected Snellen static visual acuity of each participant. The illuminance at the chart face was 250 lux. In addition, an estimate of each participant's contrast sensitivity and equivalent visual acuity were measured using a Vistech VCTS 6500 chart (Vistech Consultants, 1988). The illuminance on the face of the VCTS chart was 300 lux. A summary of visual screening characteristics for the two groups is contained in Table 9.

TABLE 9

Age and visual characteristics of participant groups

Age Group	Mean Age (Years)	Range (Years)	Range of Snellen Static Visual Acuity	Range of Equivalent Visual Acuity Using Vistech VCTS
Young	23.67	21-28	6/3.9 - 6/7.5 (20/13 - 20/25)	6/4.5 - 6/7.5 (20/15 - 20/25)
Older	73.17	65-82	6/6 - 6/15 (20/20 - 20/50)	6/6 - 6/21 (20/20 - 20/70)

Vehicle and Test Area

A 1989 Honda Accord was instrumented with a fifth wheel measuring device. The fifth wheel was mounted to the rear of the car by a tow bar. Instrumentation wiring from the fifth wheel was fed through the trunk to a digital readout box, mounted to a platform fitting between the two front seats. The distance from the center of the fifth wheel to the driver's eye was 3.5 meters (11.5 feet). This correction factor was taken into account during data transcription.

Based upon the results of Pilot Study III, site 1 (see Figure 3) was chosen to be the only one where signs were displayed. A diagram of the test track area is shown in Figure 13. The distance from the calibration start point to the target warning symbol sign was 579 meters (1900 feet). Researching previous legibility distance data, the distance available of 579 meters was enough to avoid a ceiling effect. All three pilot studies also confirmed that distance to be sufficient.

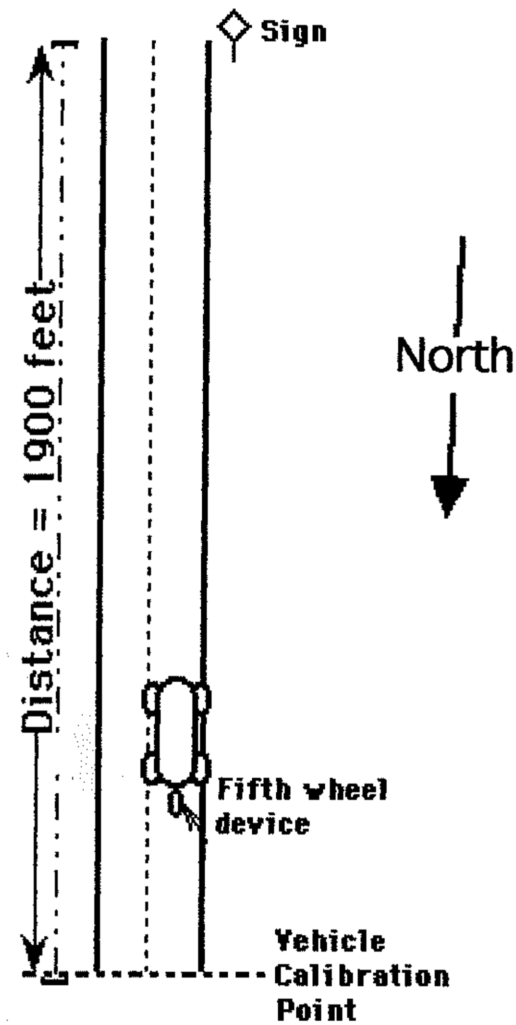

Figure 13. *Schematic of test track area (not to scale)*

Warning symbol signs were mounted on a swing-away, quick change wooden pole, specially constructed for this study. Figures 6 and 7, above, show the pole and its design to hold warning signs. As described above, the distance from the pavement to the bottom of the sign was 2.1 meters (7 feet) The pole was anchored so signs had a lateral placement of 1.83 meters (6 feet) from the edge of the traveled way, (in accordance with the *Manual on Uniform Traffic Control Devices*, 1988).

SIGN STIMULI DESCRIPTIONS

Based upon the signs used by Paniati (1988), Zwahlen et al. (1991) and Bowen (1993), six warning symbol signs were chosen for investigation in both the field and laboratory studies. The same signs as those investigated in the studies referenced above were used so legibility distances could be compared. Figure 14 shows the six symbol signs used.

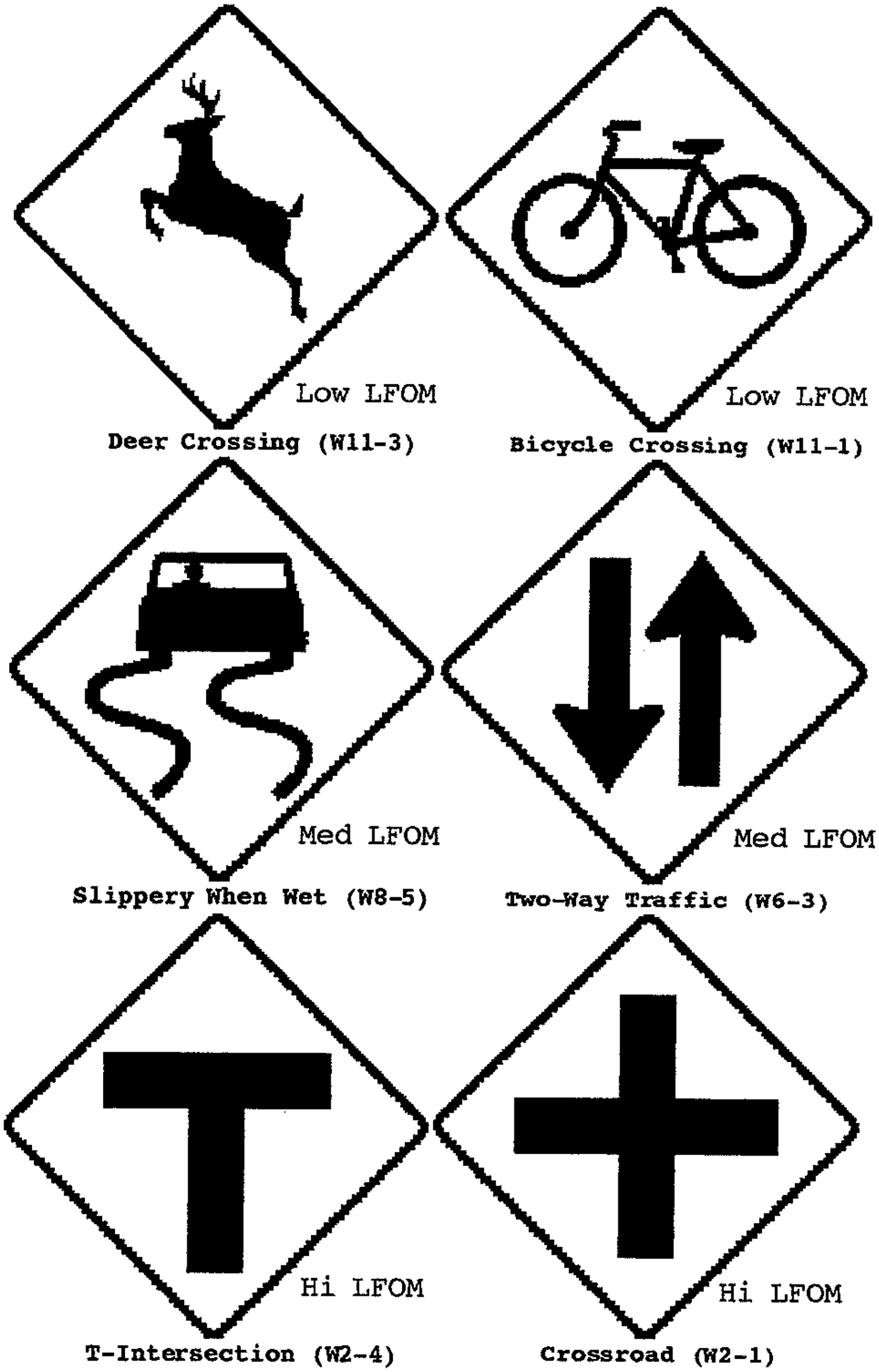

Figure 14. *Six warning symbol signs of primary interest in field and laboratory studies*

The original choice of symbol signs was also based upon the signs' legibility figure of merit (LFOM), a concept developed and tested by Bowen (1993). LFOM is related to the spatial frequency content of the symbol sign and is indexed against age and contrast sensitivity curves (Owsley, Sekuler and Siemsen, 1983). Two signs each representing low, medium and high LFOM were selected. For the low LFOM, the two signs were *Bicycle Crossing* and *Deer Crossing*. Medium LFOM values were represented by the *Two-Way Traffic* and *Slippery When Wet* warning signs. The high LFOM signs were *T-Junction* and *Crossroad*. Figure 14 labels the six primary signs by their LFOM rating - low, medium and high LFOM. The LFOM concept was not investigated further, based upon Bowen's results. Bowen found that LFOM did not adequately predict legibility distances for a range of ages of subjects.

Twelve other warning symbol signs served as distractors. The twelve signs used were:

a. *Divided Highway Ends* (W6-2)
b. *Hill sign* (W7-1)
c. *Right Turn* (W1-1R)
d. *Left Turn* (W1-1L)
e. *Reverse Curve Right* (W1-4R)
f. *Reverse Curve Left* (W1-4L0
g. *Winding Road Right* (W1-5R)
h. *Winding Road Left* (W1-5L)
i. *Side Road Right 90 Degrees* (W2-2)
j. *Side Road Right 45 Degrees* (W2-3)
k. *Y-Intersection* (W2-5)
l. *Right Lane Ends Merge Left* (W4-2)
m. *Narrow Bridge* (W5-2a)

All signs were standard 76.2 cm x 76.2 cm (30" x 30") size, with the exception of two 91.44 cm x 91.44 cm (36" x 36") signs used as distractors: *Divided Road Ends* and *Narrow Bridge*. Every warning sign was diamond shaped, with a yellow background and black symbol and border. Signs were made from new engineering grade reflective material. Signs were donated to this research project by District 1, Texas Department of Transportation, Paris, Texas.

A research assistant was positioned close to the sign site and responsible for changing signs in a timely fashion.

PROCEDURE

Research participants were met at Riverside Campus, Texas A&M University. Each was driven to a building where he/she was briefed on field study procedures and measures of visual performance were taken. Special consideration was given to explaining the precise nature of response required to the warning symbol signs during this training period. The instructions given to the subjects were as follows:

Instructions to Participants: Field Study

"You are participating in a study which will determine the best method to predict distances at which you can see and correctly identify warning symbol signs. There are two phases to this study and you are participating in Phase I today. You must also be available for another 1 and 1/2 to 2 hour period for Phase II. Phase II will be a laboratory setting and will occur in several weeks.

Phase I involves you driving my car along a marked test track area on at the Riverside Research Annex. You must maintain your speed at no greater than 25 miles per hour. This is so the signs can be changed in a timely fashion. Your task will be to maintain your speed and drive down the marked course while looking for a warning symbol sign.

There is only one site where the symbol signs will appear. Its location will never change, however, what sign is mounted at the sites will change every time you drive by. You are asked to give two responses to the sign. The first is when you feel as though you can make a good guess as to what the sign is. The second is when you are confident as to what the sign is - confident as defined as being willing to make a vehicle maneuver in response to the sign information. Also, you will make your response from a menu of 27 possible symbol signs which will always be available for your review. Simply respond by giving the sign number as it appears on the menu page. This menu is available for your easy review and referral at all times during the experiment.

There is no penalty for being incorrect in your identification of the symbol sign. You will not receive feedback on your guesses. You will only receive feedback when you give a "certain" response. If you are incorrect, you will told so and you will be asked to continue to drive the car closer to the sign to give additional responses until you are correct in the sign's identification.

The signs you will encounter will always be one of the 27 signs on the menu. However, you may not see all 27 signs and some signs you may see more than once.

Cautions to remember. Please keep your speed at no greater than 25 mph. I will remind you if you begin to accelerate beyond that speed. NEVER back up the car. NEVER use reverse gear. This is due to the "fifth wheel" measuring device mounted on the back of the vehicle. Putting the car in reverse and backing up any distance whatsoever will break the device immediately. In order to turn around, you will be making long gentle turns at the ends of the test track. I will be asking you to stop at the start point. Here you will line up a reference point on the car with a line on the pavement. This alignment of the car will allow me to accurately calibrate the measurement device.

You will not have the use of the side or rear view mirrors. This is to avoid temptation of looking in the mirrors and perhaps catching a glance at the next symbol sign being mounted as we pass by in the opposite direction.

If you have any questions at all during the experiment, please stop and ask.

There will be at least two test runs set up for you to drive for familiarization with the site location and the task which I am asking you to perform. We will do those two practice runs. If you feel you need or want any more practice, please indicate thusly and another practice trial will be ready for you. I only want you to begin the actual data collection trials when you are thoroughly comfortable with the task of driving my vehicle, seeing where the signs will be located and the two responses I am asking you to give. Do you have any questions?"

Training of Participants

Additionally, training was applied to ensure familiarity with and identification of all potential 27 warning symbol signs. Figure 15 is the pictorial menu of 27 signs presented to the participant for training and used by him/her during the course of trial runs i.e., this same menu was available to each subject at all times while driving the research vehicle. As much time was used as required by a participant to review and feel comfortable with the menu of signs he/she would possibly encounter.

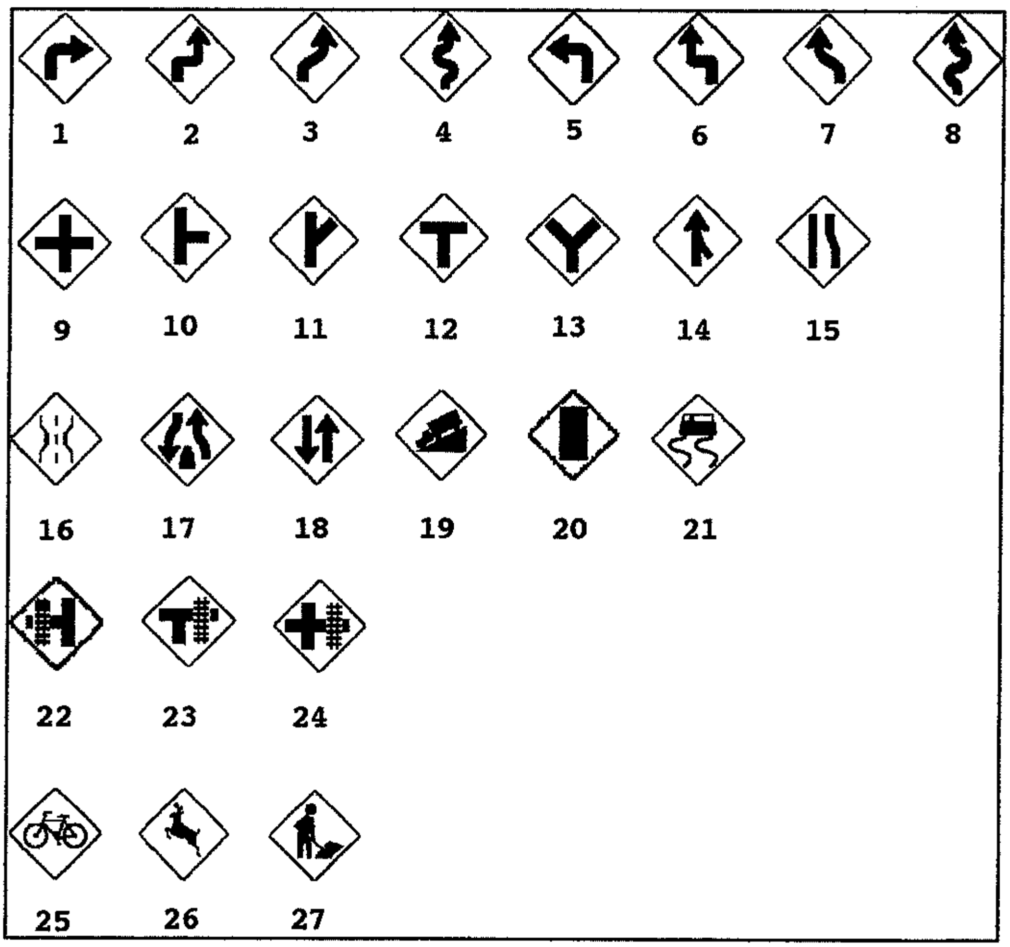

Figure 15. *Menu of 27 possible warning symbol signs*

Each participant spent time looking over the menu and asking about a sign which they did not recognize or were uncertain about its meaning. Although subjects only had to respond by the menu number of the sign they encountered, the subjects felt it was important they were familiar with all signs to make the response easier.

The older subjects remarked specifically the need they felt to be 100 percent familiar with all symbol signs and have a title or name attached to call each one, even though only a number was required. Perhaps due to their unfamiliarity with some of the symbol signs, the menu presented a challenge to learn the names of symbol signs they might encounter in their future driving experience.

In Figure 15 it can be seen, to the extent possible, the signs were grouped by classification in lines on the menu. The first line shows the eight signs which depict turns and curvature in the roadway. The first four in the top line represent curvature in the road which initially involves a right turn or curve. The second four are mirror-images of the first four, each showing the same sign, but with an orientation to the left. The fourth row are all railroad intersections with roadways. The fifth row signs fall into a "miscellaneous" category.

After the subjects' examination of the menu, the experimenter reviewed the menu layout and how signs were grouped into categories by line on the menu. When no questions were remaining and the subjects said they felt comfortable with the signs to be observed, experimental trials commenced.

Subjects were informed that the signs encountered during the experimental session would always be one of the 27 signs on the menu. However, they were also told they may not see all 27 signs, and some signs might be seen more than once.

Test Procedure

The participant then drove the research vehicle to the test area. Subjects always drove the car during all trials so as to add realism to the field study scenario. Subjects were instructed how to properly align the car for calibration of the fifth wheel measuring device. A line painted on the pavement was visually lined up with the lock on the driver's door. At this point, the digital readout was zeroed for the fifth wheel. This was accomplished before the beginning of each trial.

Response Variables

Subjects were instructed to make two responses. The first was when they could make a guess as to the sign's identity. The second, or a "certain" response, was when they were confident as to the sign's identity. "Confident" was defined as being willing to make a vehicle speed maneuver in response to the sign information (*Manual on Uniform Traffic Control Devices*, 1988).

The first type of response, a guess, if correct, was termed "recognition distance." If the response was a correct guess, no feedback was given, but the sign guessed by number and recognition distance were recorded. If the guess was incorrect, the response and distance were also recorded but the data not analyzed. Future analyses of guessing patterns leading to a correct response can be performed. Although it was the intention of the author to not have specific feedback as to correct or incorrect guessing, the fact that no response was given was, in actual fact, positive feedback. The failure for the experimenter to respond to an incorrect guess was feedback.

The second response, a certain identification, was for the purposes of this study termed sign "legibility distance." The criterion for a certain response is given above. Once the correct response was given, the subject was requested to turn the car around and return to the start point. At that time, the car was visually recalibrated with the line across the roadway and the fifth wheel rezeroed. If the response prefaced with "certain" was incorrect, the subjects were informed their choice of match with one of the 27 signs was incorrect. They were instructed to continue driving toward the sign, giving further guesses or certain responses until the sign was correctly identified. As noted above for guesses or "recognition distance," the subjects were indeed given feedback to an incorrect certain response. Feedback was not controlled for as well as was intended.

Subjects made their responses from the menu of 27 possible signs by responding with the number of the sign which matched their perception. This response technique is the same used by Bowen (1993) in his laboratory study. This response procedure controlled for learning of the signs by not having to memorize a meaning or name assigned to each sign image. The subjects had to match their perception with the number of the sign on the menu. No description of the symbol was required, as in previous studies. Therefore, the variance in responses would be expected to be minimized because it was a simpler response requirement than symbol description.

A minimum of two practice runs were performed before the data collection trials began. For the practice runs, two signs were displayed which were not ones of primary interest.

During practice and data collection trials, subjects accelerated from the start point to a speed not to exceed 40 km/h (25 mph) toward the symbol sign which was 579 meters (1900 feet) away. Subjects were instructed to monitor and maintain their speed at or below 40 km/h during all trials.

Apparatus

A Panasonic VHS video camera, mounted on a tripod stabilized in the back seat of the vehicle, videotaped the digital readout of the fifth wheel. The camera's audio channel was also used to record the subject's responses. Each videotape was transcribed after each data collection session to record legibility distances and responses.

The experimenter sat in the back seat to monitor the video tape, zero the fifth wheel digital readout, answer any questions from the subject and provide feedback when a "certain" response was given.

Experimental Design

A repeated measures design was used in this research, whereby the 24 subjects served in both the field and laboratory studies. Between subject factors were age, gender and subject nested within age and gender. The within subject factors were individual sign (*Cross Road, T-Junction, Slippery When Wet, Two-Way Traffic, Bicycle Crossing and Deer Crossing*) and condition (either field or laboratory).

Six replications of the six signs were collected from each subject. A description of the sensitivity analysis performed on Pilot Study III's ten legibility distance trials is found above.

A total of 60 trials were collected on each subject - six replications on each of the six signs of interest, while the other 24 were for 12 distractor signs. A random order for the 60 trials was constructed using a table of random numbers. This random order for the six replications of six target signs, along with the distractor signs, was the same for all subjects. The reason for using the same random order for all subjects was driven by several factors. First, as mentioned above, the session length would have been excessive due to the time involved with handling 18 different signs in a random order each time. The extra amount of time involved would have caused the subject to become fatigued. The fatigue may have affected his/her responses and confounded the data. Also the sun angle was being controlled by running sessions during certain times of the day. That control would have been lost if the experimental sessions were longer than the two to two and one-half hours required to finish the random order of 60 trials. It obviously would have been preferable to administer a different random order to each subject.

The briefing, visual screening, training and running of 60 trials defined the experimental session which required between two and two and one-half hours.

Dependent Variables

The two types of responses briefed to the subject are explained above. Recognition distance was that distance at which the first correct guess of the symbol sign was made. Legibility distance was the distance from the sign where the subjects was certain as to the sign's identification and the response was correct. Both of these dependent variables were collected from the digital readout of the fifth wheel. The distance on the readout was subtracted from 579 meters (1900 feet) to obtain the distances, expressed in feet.

Luminance Measurements

All trials were collected in the months of July and August 1993, during daylight hours between 8:30 a.m. and 4:00 p.m. Sign luminance level was measured three times during the day: before the first subject, generally at 8:00 a.m., mid-day, and at the end of the day, approximately 4:15 p.m. A *Left Turn* sign was erected at the site, as it allowed for maximum yellow background area to be sampled in four places: top, right side, bottom and left side. Five luminance samples were taken at each of the four locations on the sign. Luminance measurements were made with a Tektronix Model J16 digital photometer, which was equipped with a Model J6531 1° narrow angle luminance probe. Table 10 shows the result of a typical set of luminance measurements.

TABLE 10

Sign luminance measurements made on 7 July 1993, clear day, *Left Turn* sign

Sign Area Sampled	**Morning Sample 7:50 a.m.**	**Mid-Day Sample 11:30 a.m.**	**Afternoon Sample 4:20 p.m.**
Left	1317.6 cd/m^2	1449.9 cd/m^2	2002.1 cd/m^2
Top	1326.5 cd/m^2	1543.1 cd/m^2	1893.2 cd/m^2
Right	1377.9 cd/m^2	1544.4 cd/m^2	1938.4 cd/m^2
Bottom	1483.5 cd/m^2	1562.9 cd/m^2	1890.5 cd/m^2

All experimental sessions were conducted on days where the weather was either clear or partly cloudy. On clear days, there was no haze. There were always sunny conditions prevailing during all experimental sessions. Subjects did not drive using sunglasses. Only one experimental session was canceled and rescheduled due to rain.

LABORATORY STUDY

INTRODUCTION

As discussed in the section Introduction above, laboratory research was criticized by Zwahlen et al. (1991) for its lack of realism and other factors contributing to what they consider to be relatively short recognition distances when compared to field data.

They described past laboratory studies as lacking in approximations of the real world "...in terms of resolution, luminances, colors and contrast values." In addition, Zwahlen et al. conclude that "...studies conducted in laboratories using monochrome monitors, projectors or visual displays generated by driving simulators may not adequately represent the real world conditions." They hypothesize that "...laboratory studies produce much smaller symbol recognition distances than the field studies. The difference in recognition distance could be due to factors such as insufficient display resolution, insufficient luminance and contrast representation, a dynamically changing (increasing in size) sign and symbol while the actual viewing distance of the subject to the display stays always the same (no change in depth perception), small image vibrations, non-uniform and less sharp symbol or legend contours, etc."

The laboratory portion of the present research reported here was designed first of all to answer the criticisms brought forth by Zwahlen et al. To the extent possible, the factors Zwahlen et al. hypothesize as contributing to legibility distance differences between laboratory and field experimental settings, or from one laboratory to another, were optimized or controlled.

As a second purpose, this study was aimed at examining field and laboratory-generated legibility distances using the same subjects. The expectation, in terms of the null hypothesis, was that the repeated measures design would have subjects giving the same performance under both field and laboratory conditions. Heretofore, no research had systematically varied the same factors in both a field and laboratory setting with the same subjects providing multiple observations. In all previous studies researched, the experimenter chose either a laboratory-based or field-based setting. The overwhelming selection of setting was a laboratory-based one.

Replications and the study of within subject variability had not been dealt with in the literature researched for this study. Only one study was found, Zwahlen et al., 1991, where replications were used. In the Zwahlen study, only two replications were made on the 12 signs for which they were gathering legibility distances. However, all the subjects in Zwahlen's research were young drivers. No older drivers were used in either his daytime or nighttime legibility study. From Pilot Studies I through III, the within subject variability was found to be present and proved to be an interesting topic to examine in depth in both the field and laboratory studies.

In his unpublished Master's thesis, Bowen (1993) collected six replications on each of the six signs under study. Bowen provided raw data from which coefficients of variation could be computed for his subjects (Charles K. Bowen, personal communication, March 1, 1994). Again, one drawback with generalizing from Bowen's data is the narrow age range of his 15 subjects. The mean age of subjects in Bowen's study was 32.8 years, with a range of 22 to 56 years.

Thus, like the field study summarized above, the goal of the laboratory study was to identify important factors and their interactions affecting warning symbol sign legibility distance. Instead of relating studies by different authors who used distinct experimental methods and different subjects, a repeated measures design allowed for direct comparisons between field and laboratory data with the same subjects. The comparisons made could provide a basis for the development of reliable evaluation techniques for new signs, suitable for implementation by traffic engineers. The laboratory portion of this research involved a lengthy development of a simulation technique that addressed shortcomings found in the literature. This technique is outlined in detail in this section. Appendix A includes a very detailed description of the software used, lessons learned and developmental steps for the reader who is interested in the software development side of the simulation technique.

METHODOLOGY

Research Participants

The same 24 subjects who participated in the field study took part in the laboratory portion of this research. A check of any vision changes occurring over the 11 weeks since the field runs occurred revealed none.

Apparatus

A Macintosh IIci computer and Viewsonic® 7, 43.18 cm (17 inch) diagonal high resolution color monitor were used to present the laboratory test objects. The Macintosh IIci computer had 8 MB of RAM storage and a 100 MB hard drive. Table 11 summarizes the specifications of the Viewsonic® 7 FS Multi-Frequency color monitor.

TABLE 11

Specifications of Viewsonic® 7 monitor used in laboratory study

Input Signal - Video signaling system	RGB analog
Cathode Ray Tube (CRT) Size	43.18 cm diagonal (17 inches)
CRT Dot Pitch	0.28 mm
CRT Phosphor	RGB short persistence
CRT Surface Treatment	Silica Coat ESF (Electro-Static Field) Treatment

An NEC Technologies, Inc. MacFG™ 8X,24X Color Display Interface card and RadiusWare™ software were used to drive the Viewsonic® 7. Specifications on the interface card are found in Table 12.

TABLE 12

Specifications for MacFG™ 8X,24X Color Display Interface card

Screen Size Setting for 43.18 cm (17") monitor	MultiSync5FG
Resolution	72 dots per inch 896 x 672 lines resolution
Flicker-free refresh rate	76 Hz
Number of colors displayed	2^8 or 256 colors
QuickDraw Acceleration	Yes

The QuickDraw acceleration feature in Table 12 above dramatically improved screen performance when working with color-graphics applications like the ones used in the laboratory simulation to be described later.

Experimental Setting

Subjects were seated behind a table 6.1 meters (20 feet) away from the computer monitor. Figure 16 is a diagram of the laboratory experimental set-up. The experimenter sat at a table adjacent to the monitor and used a template on the keyboard to input subjects' responses. A photograph of the keyboard and affixed template is shown in Figure 17.

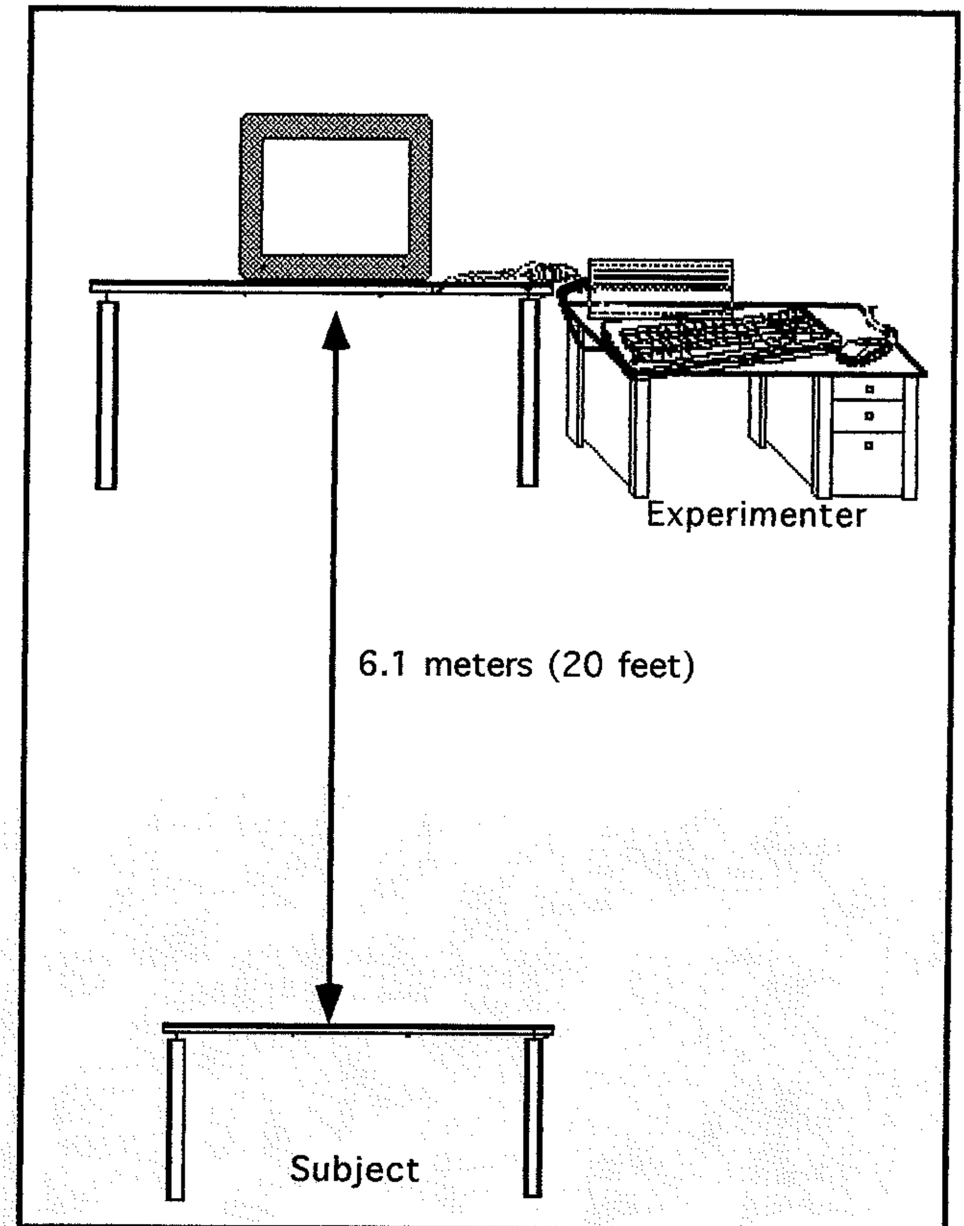

Figure 16. *Schematic of experimental setting*

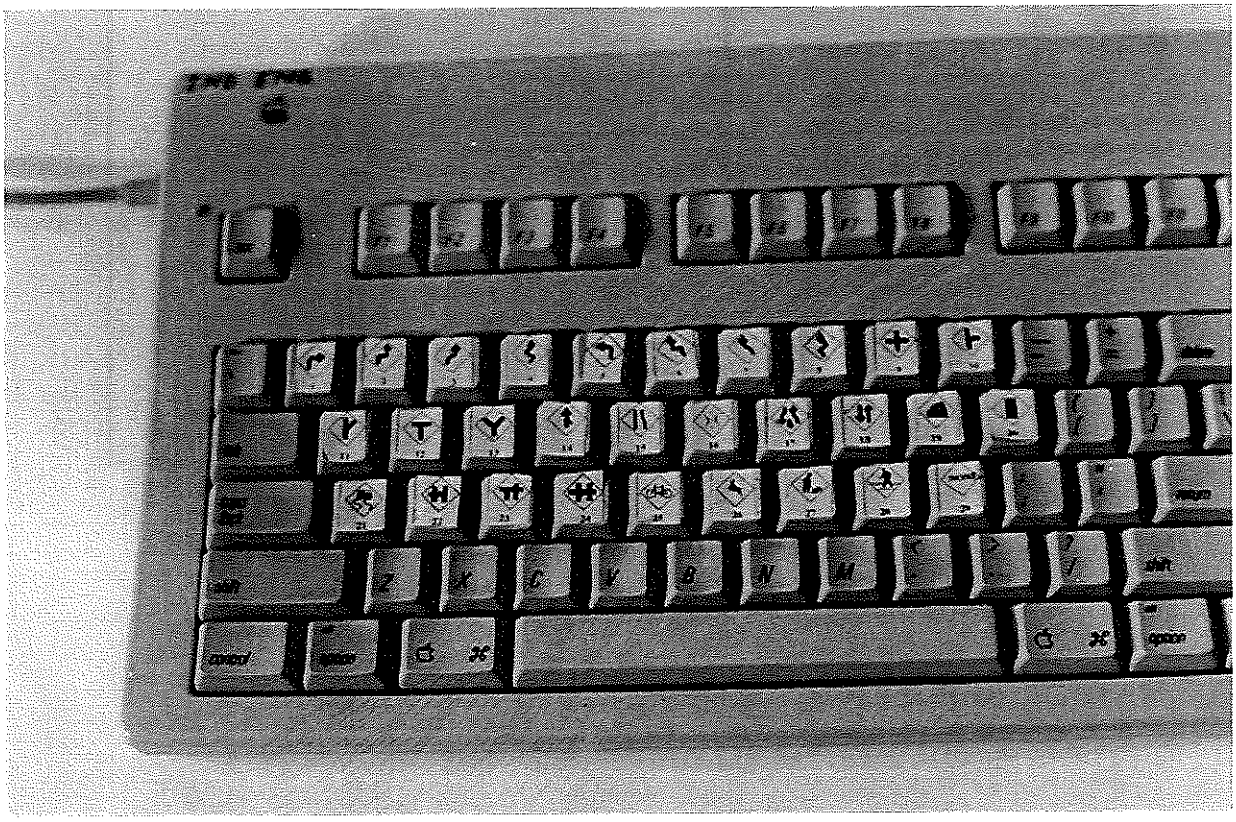

Figure 17. *Photo of keyboard template for response input*

ANIMATION/SIMULATION OVERVIEW

In an attempt to provide a high fidelity simulation of the driving environment, this experiment used a high-resolution color monitor and software to generate an animation of the real world. The computer and monitor are described above.

Software packages were used to capture and digitize a video image of a background scene of the Riverside Campus very similar to that encountered in the field study. Also, graphics software was used to generate the warning symbol signs and size them appropriately. Animation software was used to place the correct size symbols in the background to create the simulation of moving toward the sign, just as the subject did in the field study. Additional software controlled the experiment, collected and stored the data.

Early Simulation Design Development

Since the original goal was to develop a design tool for use by district engineers designing new traffic signs, a low-cost prototype was a priority. With simplicity of equipment and readily available software requirements of foremost concern, a laboratory simulation was built around these specifications.

Background Capture

In order to have the background image the same as subjects encountered in the field study, a number of methods were tried to capture and input an image into the computer. To that end, the following techniques were tried, with the results summarized.

Option 1

VHS Video capture: A Panasonic Model AG-180, Professional/Industrial Video VHS Reporter camera was used to film the treed background at Riverside Campus. The camera was connected to a Macintosh Quadra 700 computer which contained a RasterOps® 24STV graphics display board. Video-capture software, MediaGrabber™, Version 2.0, captured and digitized the image into a 24 bit format. The resulting image was too grainy, due to the VHS low-resolution format. In search of a higher quality video image, the next alternatives were explored.

Option 2

Betacam SP Camcorder: This camera proved to be a very expensive option. To use this camcorder option to capture a background image for this project would entail hiring a video specialist to shoot and capture the image with camera and tape at a cost of over $150/hour. Labor, location shooting costs, plus costs of video tape would raise the full lease of the camera to over $200/hour. An additional cost of having the Betacam image captured and saved in a tag-based image file format (TIFF) file would be $60 per hour. This TIFF format would then have to be converted to a Macintosh-readable file. Total lease expenses would amount to approximately $250 to produce one digitized background image. This expenditure was determined to be too high for the low-cost prototype goals.

In order for a Department of Transportation (DOT) to use this method to record background scenes would require the purchase of the SP camcorder at an investment of approximately $15,000. A video production facility is also necessary to take the camera tape and play it through editing equipment. Unless many background images would be required, the investment in camera and facility, with personnel to staff it, would not be cost-effective.

Option 3

Professional Hi8 with three Charged-Couple Devices (CCDs) Camcorder: This possibility was also more costly than objectives permitted. The Hi8 image would have been captured and then saved as a TIFF file using the same procedure as in 2. above. The TIFF file, in an MS-DOS format, would then have to be converted using Apple FileExchange© software into a Macintosh-compatible format. This Macintosh file could then be digitized as a PICT file on the Quadra 700 computer with the same graphics display board. Precise costs are not available, however, they are expected to exceed the $250/image estimate in 2., since the lease of the camera is more costly. For a DOT to purchase a professional Hi8 camcorder would be $7800. The same video production facility is necessary, as described in Option 2 above.

Option 4

Sony Prosumer V5000 Hi8 camcorder: This camera was accessible and affordable, being within the development cost-constraints of a design tool for Transportation Districts. The advantage of using this choice was that the image could be captured directly into the Macintosh Quadra 700, instead of having to undergo the intermediate step of a TIFF file format conversion as both options 2. and 3. above required. Costs were reasonable, with labor and tape expenses at less than $50/hour. If DOT wanted to purchase this camcorder, the investment cost would amount to approximately $4,000. No video production facility is required.

Option 5

Other Hi8 or Super-VHS Consumer Video Cameras: These cameras are commonplace consumer items. Yet preliminary tests showed that the image quality did not exceed the original VHS format described in 1. above. Test video of the same background image was taken using on a Panasonic Model AG-450 Super-VHS camera and a Panasonic Hi8 camera. The resulting digitized image capture was nearly as grainy as the VHS image format. The resolution, digitized in a 24-bit format allowing for 2^{24} colors, was unacceptable.

Based upon a combination of quality of image and cost to produce, option 4. above was chosen. A tripod was used to secure the camera in the roadway such that the image viewed through the camera was positioned to simulate driver's eye height and position. Several scenes were shot, some with the zoom capability of the camera.

This camera with video tape was then connected to the Macintosh Quadra 700. MediaGrabber™, Version 2.0 software captured a frame from the video tape. This frame capture was digitized and stored at 24 bit resolution as a PICT file image. The quality of the background image from the Sony Hi8 camcorder displayed in eight-bit resolution on the Viewsonic® 7 monitor was adequate. There was much less loss of color or image graininess (pixelation), as resulted with the other techniques of video capture described above.

Adobe Photoshop™, Version 2.5 software was used to smooth out any imperfections in the video capture and enhance the resolution to the maximum extent possible For instance, there were considerable amounts of grass and weeds growing up through the cement cracks in the road surface at the time of the video capture. This grass was "erased" by a smoothing technique in Adobe Photoshop™ which blended the grass out of the image. Pixels could be colored or lines made distinct in the video image with the Photoshop™ software. Figures 22 and 23 are color photographs of the computer screen displaying the background and symbol sign. Figure 18 has the *Crossroad* sign embedded in the background, Figure 19, the *Bicycle Crossing* sign.

Stimuli Development

An animation of the sign in the background was required to simulate driving toward the warning sign. This animation was performed with the use of MacroMind Director®, Version 3.1.1. This software achieves its animation by placing "cast members" on a "stage." There is an accompanying script which changes the "cast members," and controls the speed of the animation and the interaction with the program.

To accomplish the animation, cast members of signs had to be created. Appendix B summarizes in detail the procedure used for this research project to produce and edit the symbol images.

Figure 18. *Actual computer screen displaying background capture and Crossroad sign*

Figure 19. *Actual computer screen displaying background capture and Bicycle Crossing sign*

Visual Angle and Sign Image Sizing

All real-world warning symbols signs of concern were a standard size 76.2 cm x 76.2 cm (30 inches x 30 inches), in accordance with the Standard Symbol Signs (SSS) Manual. A calculation of visual angle subtended by the sign at distances from 579 meters (1900 feet) to 0 meters was performed. The formula used was:

Visual Angle (degrees) =

$$(\text{Integer}(((1000 * \text{Arctan}(((30 * \sqrt{2}) \div (\text{Distance from Sign in Feet} * 12)))) + 0.50)) \div 1000)$$

This formula comes from the determination of visual angle, whereby:

$$\text{Visual Angle (degrees)} = \text{Arctan (Size of Sign} \div \text{Distance)}$$

Calculations and plots of visual angle were done with CA-CricketGraph® III software where sign and distance measurements were input in feet. Figure 20 is a plot of visual angle subtended as a function of distance from a 76.2 cm (30 inch) sign.

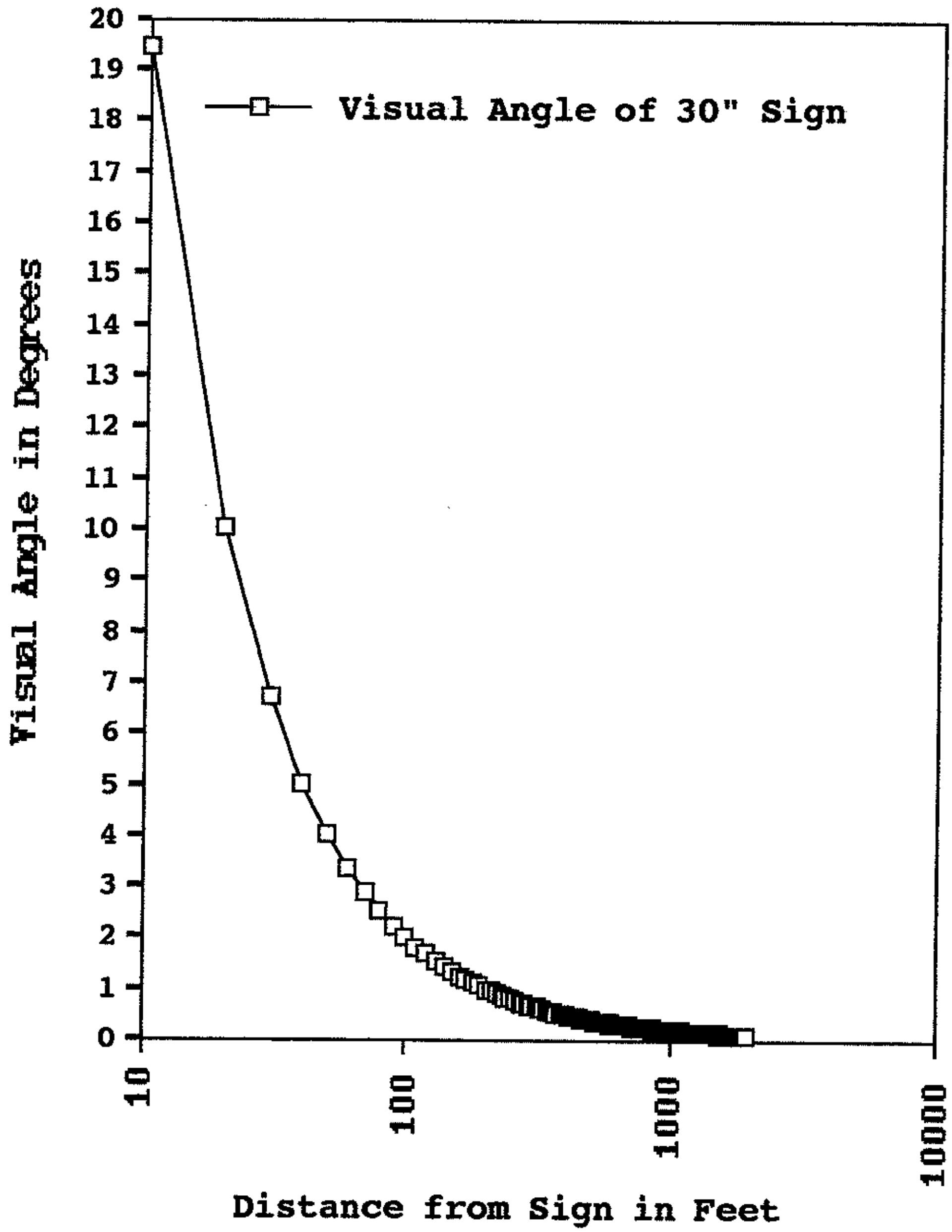

Figure 20. *Log plot of visual angle in degrees subtended as function of distance from 76.2 cm (30") sign*

Knowing the 256 x 256 pixel sign image had to be scaled down in size to be presented on the 43.18 cm (17 inch) diagonal computer monitor, a tradeoff analysis was performed to answer two questions: (1) at what distance from the monitor should subjects view the simulation? (2) what is the size of the scaled image required to match that visual angle of the sign embedded in the background to simulate viewing from the extremes of 579 meters (1900) to 0 meters away from the sign? Three viewing distances from the monitor were selected for analysis: 3.05, 6.1 and 9.15 meters (10, 20 and 30 feet).

Based upon the visual angle subtended by the 76.2 x 76.2 cm sign, the size of the sign when viewed from the three distances was calculated. It was discovered that if subjects were seated at 3.05 meters (10 feet) from the monitor, the physical size of the sign on the monitor would be too small at the longer distances. For example at 579 meters (1900 feet), the size (diagonal) of the sign image would need to be 0.57 cm (0.224 inches). With display resolution being 72 dots per inch, the resulting sign image viewed from 579 meters would consist of only 16 pixels.

The size from a viewing distance of 9.15 meters (30 feet) had the opposite problem at closer distances to the sign. For instance, at 91.44 meters (300 feet) from the sign, the sign image size on the monitor would be 10.8 cm (4.24 inches). The large size of this image would obliterate the majority of the background scene.

Therefore, seating the subject at 6.1 meters (20 feet) from the monitor was determined to be a satisfactory alternative. Sign sizes in order to subtend the correct angle from distances approaching the sign were computed in 3.05 (10 foot) increments. These sizes were translated into pixel diameter of the sign image.

Based upon the results of the field study, it was decided the entire distance from 579 to 0 meters did not need to be represented in the simulation. The worst case sign, *Bicycle Crossing*, had the closest legibility distance. That distance of 57.9 meters (190 feet) was used as a terminal point, with another 9.1 meters (30 feet) subtracted to give a comfortable latitude for sign legibility for the simulation.

Animation Particulars and Peculiarities

MacroMind Director®, Version 3.3.1 by MacroMedia© operates by placing "cast members" on the stage. The discussion of the process for creating scaled sign image cast members is outlined in Appendix A. Two cast members were required to represent each sign for each simulated viewing distance - the diamond and the symbol sign element.

The digitized background scene was always present and did not move during the animation by MacroMind Director. The yellow warning diamond was placed in the background scene, with calculations for its placement location accomplished based upon the following:

a. the size of the sign (in terms of pixels which translated into visual angle) from 579 meters (1900 feet) to 160 meters.
b. The lateral displacement of the sign from the roadway.
c. The height of the pole (2.1 meters or 7 feet) upon which the sign was mounted.
d. The simulated approach speed of toward the sign.

Figure 20 shows that visual angle subtended by the sign at the longer distances does not change much. This fact, combined with the integer pixel rounding peculiarity of Canvas™ (see Appendix A), translated into fewer symbol sign elements than the 165 steps between 579 (1900 feet) to 160 meters, one for every 3.1 meters (10 feet). The total number of cast members required for this animation was 87.

The speed of animation was chosen to be three frames per second. Since every frame stood for a change in distance of 3.05 meters (10 feet), the resultant simulated approach speed was approximately 33 km/h (20.5 mph). This approach speed was only slightly slower than the field study, where subjects were urged to remain at or below 40.2 km/h (25 mph).

A script was built, which consisted of frames being occupied by cast members. The frames "being played" by MacroMind Director® is analogous to frames playing in a movie film, hence the animation. The frames in each sign animation were composed of the same background scene, the diamond and the symbol element. Each frame corresponded to viewing the symbol sign from a specific distance.

Through the use of a "script" facility in MacroMind Director®, control of the animation was possible. Special scripting, similar to HyperTalk® language, was used to not only control the animation, but send commands to collect and store the sign and legibility distance responses. HyperTalk® is a script language which enables the user to write English-like statements which respond to events. In the case of this simulation, the event was the experimenter pressing one of the template keys on the keyboard.

The key pressed recorded the sign with which the subject responded, the simulated legibility distance in feet when a response was given and whether or not the response correctly matched the sign in the animation. If the response did not match, the simulation continued. On the other hand, if a match was made, the animation terminated and the next animation sequence was loaded in the computer. A file was written which contained the subject number, each sign response, simulated legibility distance from the sign and a minus sign if incorrect, and a plus sign if correct.

PROCEDURE

Research participants were met in the Human Factors and Ergonomics laboratory at Texas A&M University. As in the field study, it was felt that extra consideration must be given to elucidate the nature of response required to the warning symbol signs subjects would see in this phase of the research. During a training period, (see Appendix B for instructions given to participants) instruction was given to ensure familiarity with and identification of all potential 27 warning symbol signs. Figure 21 is the pictorial menu of the 27 signs briefed to the participant during training and used by him/her during the course of the trials in the laboratory session (same as Figure 15).

This menu was made available to each subject at all times during the laboratory experiment. Exactly as in the field portion of this research, as much time as required and requested by a participant to review the menu of signs was allowed.

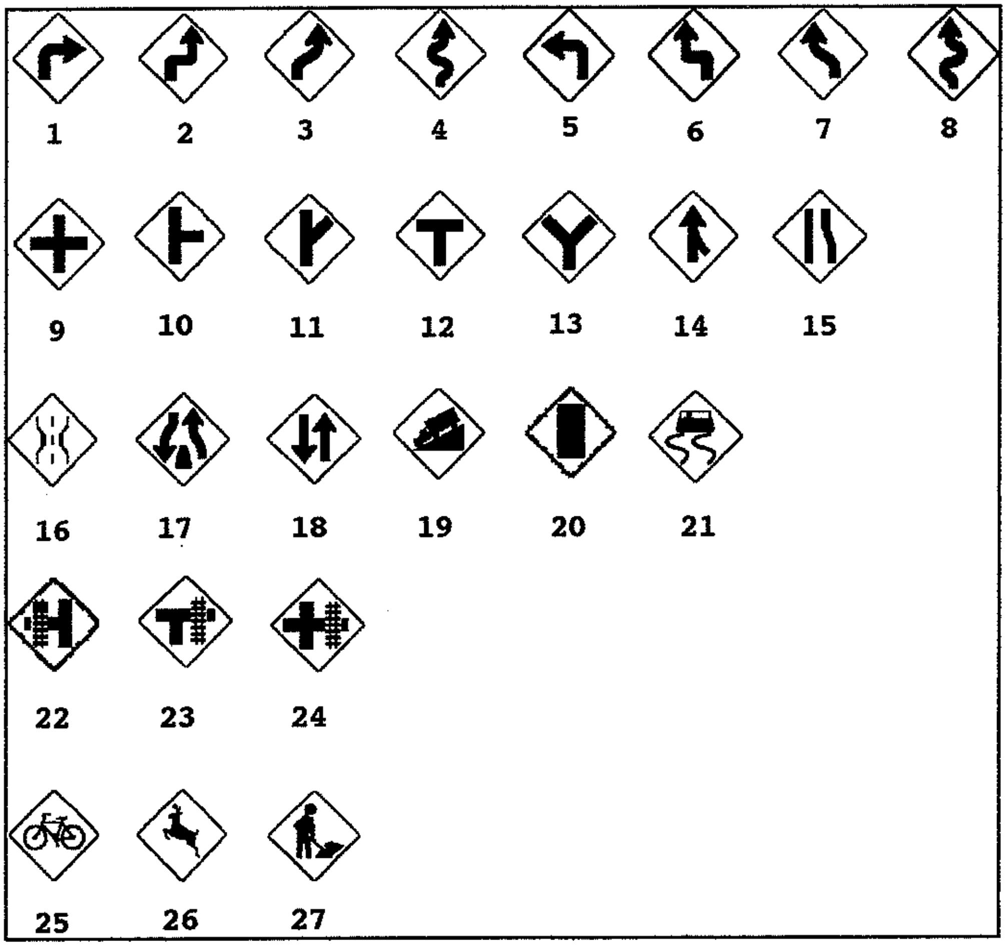

Figure 21. *Menu of 27 possible warning symbol signs (repeated)*

Because each participant had already participated in the field study, he/she was accustomed to much of the instruction and training. The same training scheme described in above in field study was used for the laboratory study. However, between 11 and 14 weeks had transpired since the subjects participated in the field study. It was felt that a retraining period was imperative.

After the subjects' review of the menu, the experimenter went over the layout of the menu and how signs were grouped by categories by lines on the menu. Each sign was named and the subjects were given the opportunity to raise any questions. Additional review time was allowed, if requested, for subjects to study the menu. When no questions remained and the subjects said they were comfortable with the signs to be observed, laboratory experimental trials commenced.

Again, as in the field study, subjects were told the signs they would see during the experimental session would always be one of the 27 signs on the menu. However, they were also told they may not see all 27 signs and some signs might be seen more than once.

A examination and example of the responses required were given. Subjects were told they needed to make two responses - precisely the same ones as in the field study. For review, the first type of response, a guess, if correct, was termed "recognition distance." The second response, a certain identification, was for the purposes of this study termed sign "legibility distance." Subjects made their responses from the menu of 27 possible signs by responding with the number of the sign which matched their perception. A discussion of the attempt to eliminate feedback in the field study applies to the laboratory study, as well. As before, the lack of feedback became a feedback cue to the subjects.

Inasmuch as the laboratory simulation was a new experience for all participants, several practice runs were performed before the data collection trials began. The main purposes behind the first two practice sessions were to illustrate the speed of the animation, combined with the sign increasing in size behind the static background. During practice runs, at least four sign animations were constructed specifically for practice, and did not contain any symbol signs of primary interest. For the most part, after three practice runs, the subject indicated he/she was ready to start the data collection trials.

Experimental Design

As indicated above, a repeated measures design was used, whereby the 24 subjects served in both the field and laboratory studies. As before, the between subject factors were age, gender and subject nested within age and gender. The within subject factors were individual sign (*Crossroad, T-Junction, Slippery When Wet, Two-Way Traffic, Bicycle Crossing and Deer Crossing*) and condition (either field or laboratory).

Just as in the field portion of this research, six replications of the six signs were collected from each subject. A total of 60 trials were obtained from each subject; six replications on each of the six signs of interest, while the other 24 were for 12 distractor signs. As in the field study, a random order for the 60 trials was constructed using a table of random numbers, and this random order was the same for all subjects. The same random order was again chosen for all subjects. Although as discussed above, this was not the preferred procedure. However, no additional variables needed to be present which were not in the field study.

Briefing, training and conducting 60 trials composed the experimental session and required between one and a half and two hours.

Dependent Variables

Two types of responses were briefed to the subject, as explained above. For the laboratory part of this research, the same two dependent variables were collected. The distance at which the first correct guess of the symbol sign was made was termed recognition distance. The distance from the sign where the subject was certain as to the sign's identification was called legibility distance. Both of these dependent variables were collected via the keyboard entries made by the experimenter, by a process explained above.

RESULTS

LEGIBILITY DISTANCE

INTRODUCTION

Many subjects, especially the older drivers, only gave one response - a certain response. In spite of additional instructions during both sessions, these subjects would not guess sign identity for a recognition distance. Primarily for that reason, the unreliability of such data led the writer to use legibility distance, or that distance where subjects were certain as to the response. Only the legibility distance measurements were analyzed and the results are presented here.

Field Data Transcription

All data from the field study trials were transcribed from the video tapes of the digital readout counter from the fifth wheel by viewing the tapes essentially one frame at a time. The same procedure was used for recording the distance. Once the audio track verified the first correct sign identification for any particular trial, the tape was rewound a few frames. When the first utterance was made from the subject for that sign identification, the number was taken off the digital readout counter.

Laboratory Data Transcription

In the case of the laboratory studies, all responses were keyed in via template by the experimenter. If the sign identification was incorrect, no feedback was given, as in the field study, and the trial continued with the sign continuing to grow to simulate forward movement until the subject gave a correct response. If the key input matched the sign for that trial, the screen went blank and the next sign animation program was loaded in the computer. The computer was programmed via the MacroMedia Director® to write a record with the sign response, distance from the sign and the correctness of the response. Files containing records for all signs for each subject were created containing legibility distances and sign responses. As noted above for field study data, only the correct responses, the legibility distances, were analyzed.

Data Analysis Software

Data analyses were performed using JMP® Version 3, *Statistics Made Visual*™ (SAS Institute, Inc., Cary NC, 1994) a statistical software and graphics package for the Macintosh computer. The JMP® analyses used on all data uses the equivalent of PROC MIXED (SAS Institute Inc., Mainframe Release 6.08). PROC MIXED is recommended by SAS over PROC ANOVA or PROC GLM. PROC GLM has been superseded by PROC MIXED and was specifically written for repeated measures designs and random effects (Searle, Casella, and McCulloch, (1992) and McLean, Sanders, and Stroup, (1991)).

Organization of Results

Using all six replications as the response set, two data sets were created. The first was for the field data only, the second, containing laboratory data only. An ANOVA was performed on each data set separately. The first section of the Results section is devoted to the results of the field study ANOVA. The second section deals with the results found from the ANOVA of laboratory data only.

As discussed above, a large within subject variability was found for legibility distance. Since this finding was not present in the literature reviewed, priority to analysis of this variability was given. But because the within subject variability encompasses a lengthy discussion of results, an entire section is devoted to this finding. The first portion of that section examines the within subject variability found in the field data only. Analysis of the laboratory study data's within subject variability is summarized next.

Finally, subsets of the six replications of the response variable, legibility distance, are formed and means and standard deviations reported. This analysis was done to illustrate how misleading studies which included only one or at most, two replications, of the response variable can be. A summary of the study results based upon those subsets of the dependent variable conclude the section.

FIELD STUDY

An analysis of variance (ANOVA), using PROC MIXED, was performed on the mean legibility distance in order to assess the effects of age, gender, individual sign and subject and their interactions. Table 13 is the source table from that ANOVA. The analysis of variance applied is for a special case of the nested factorial design, also known as repeated measures.

As can be seen from Table 13, the main effect of gender and all interactions involving gender were not significant. For females, the average legibility distance for all signs combined was 289.1 meters (948.4 feet), while males averaged just over 3.05 meters (10 feet) closer to the sign at 285.9 meters (938.1 feet). As expected, none of the interactions involving gender, like Gender X Age, Gender X Sign, Gender X Age X Sign, were significant either.

TABLE 13

ANOVA source table for field study data only

Source	SS	MS	df	F Ratio	Prob>F
Gender	316103	316103	1	0.13	0.7263
Age	26035402	26035402	1	10.38	0.0043*
Gender*Age	4942719	4942719	1	1.97	0.1756
Sign	22465923	4493185	5	70.33	0.0000*
Gender*Sign	505435	101087	5	1.58	0.1720
Age*Sign	691344	138269	5	2.16	0.0640
Gender*Age*Sign	260031	52006.2	5	0.81	0.5424
Subj[Gender,Age]	50147055.0	2507353.0	20	39.24	0.0000
Sign*Subj[Gender,Age]	6389155.4	63891.6	100	2.49	0.0000

*** denotes significance at alpha=0.05**

Age of participants, however, was a significant main effect. Figure 22 is a plot of the main effect of age. The overall mean legibility distance, collapsed over all six signs, for older subjects was 232.64 meters (763.25 feet), and 342.11 meters (1122.4 feet) for young participants.

Table 13 above indicates the main effect of individual warning symbol sign is highly significant ($p<0.0000$). Table 14 gives the mean, standard deviation and standard error of the mean legibility distances for the six symbol warning signs investigated in this field study. Figure 23 is a plot of the six signs' mean legibility distances.

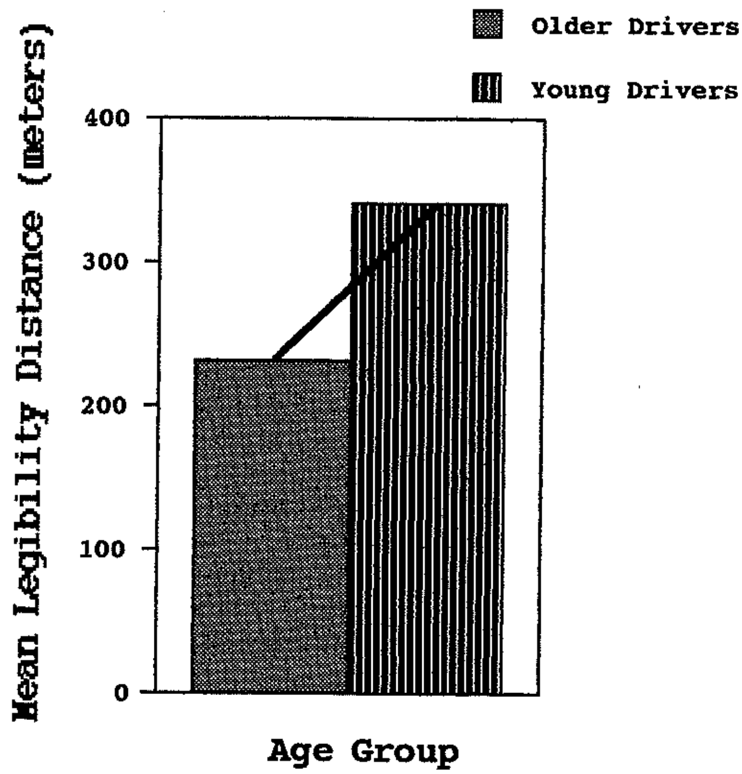

Figure 22. *Main effect plot of age for field study*

TABLE 14

Mean, standard deviation and standard error of mean for individual sign legibility distance, field study only

Warning Sign	Mean Legibility Distance *	Standard Deviation *	Standard Error of Mean *
Bicycle Crossing (W11-1)	205.01 m (672.60)	98.70 m (323.801)	8.22 m (26.983)
Slippery When Wet (W8-5)	254.35 m (834.49)	113.12 m (371.143)	9.43 m (30.929)
Deer Crossing (W11-3)	270.20 m (886.47)	112.63 m (369.532)	9.39 m (30.794)
Two-Way Traffic (W6-3)	308.27 m (1011.38)	99.10 m (325.124)	8.26 m (27.098)
Crossroad (W2-1)	336.15 m (1102.85)	104.48 m (342.794)	8.71 m (28.566)
T-Junction (W2-4)	350.25 m (1149.10)	123.30 m (404.537)	10.28 m (33.711)

*** Values in parentheses are expressed in feet**

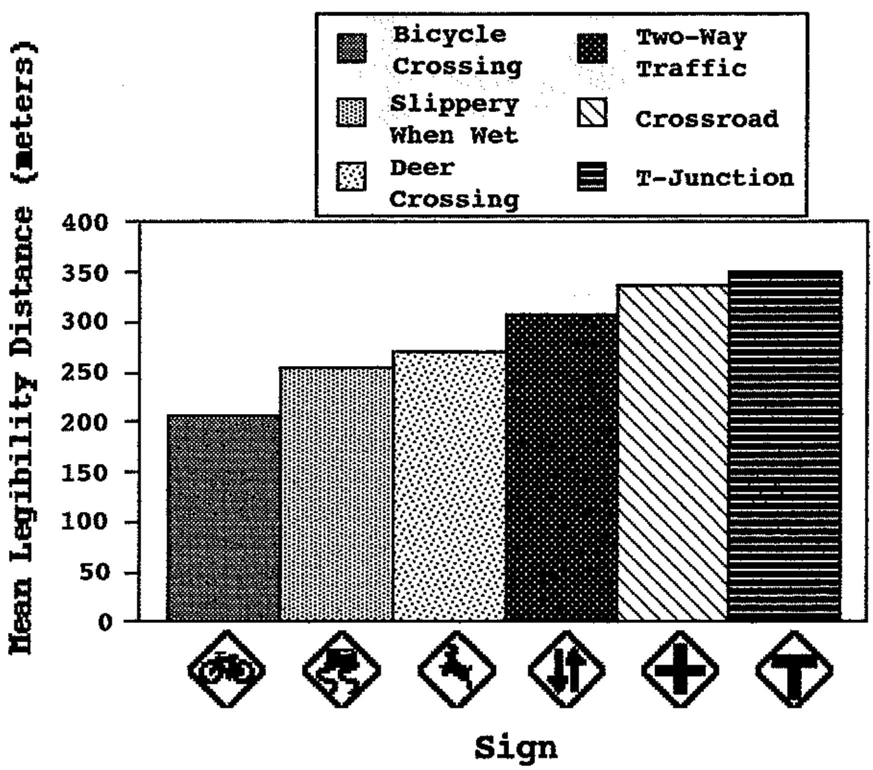

Figure 23. *Main effect plot of individual warning sign, field data*

An examination of Table 14 reveals several things. The sign with the shortest overall legibility distance (672.6 feet) is *Bicycle Crossing*. *T-Junction* has the longest legibility distance (1149.1 feet), being almost twice that of *Bicycle Crossing*. This data is collapsed over all subjects. The standard deviation is quite large (ranging between 323.8 and 404.5 feet). Table 15 shows the Coefficient of Variation (CoV) for the six signs, remembering that CoV is the standard deviation divided by the mean, expressed as a percentage.

TABLE 15

Coefficient of Variation for six symbol signs, field data only

Warning Sign	**Coefficient of Variation (%)**
Bicycle Crossing (W11-1)	48.15
Slippery When Wet (W8-5)	44.48
Deer Crossing (W11-3)	41.69
T-Junction (W2-4)	35.21
Two-Way Traffic (W6-3)	32.15
Crossroad (W2-1)	31.08

The CoV for each of the six signs, for all subjects is large. As seen in Table 15, the CoV computed upon the six replications collapsed over all subjects ranges from over 31% to over 48%. Not only does *Bicycle Crossing* have the shortest legibility distance, it also has the most variation associated with it. The least complex sign, with bold strokes, the *Crossroad* warning symbol sign, has the smallest amount of variability, as reflected by the smaller CoV. A further presentation of these results, by subject, will occur later. Figure 24 shows a plot of CoV by age group for field data.

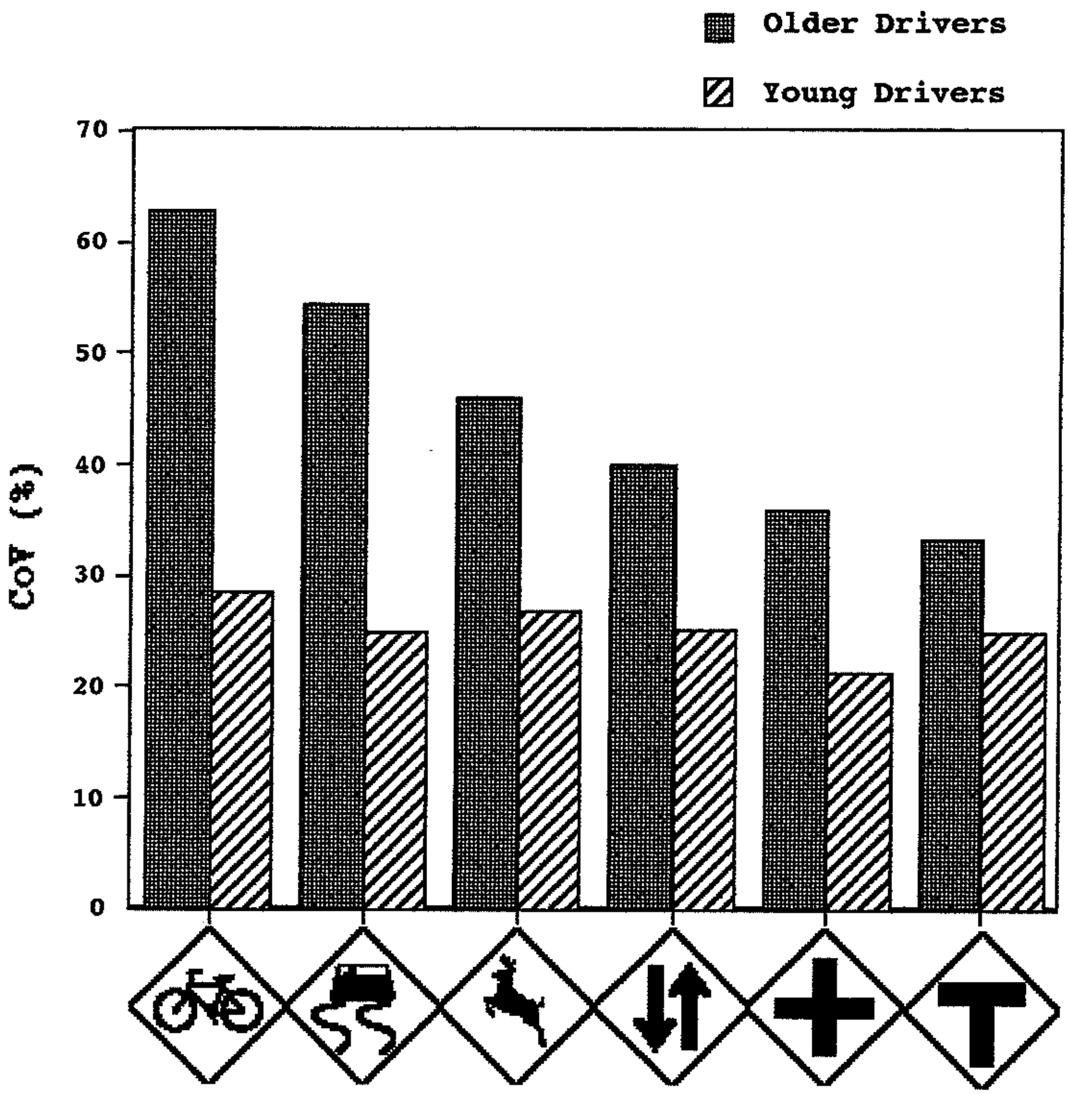

Figure 24. *Plot of CoV by age group, all signs, field data*

In all cases, for all signs, older drivers are more variable in their responses than young drivers. The difference is dramatic for *Bicycle Crossing* and *Slippery When Wet*, where the older drivers are over twice as variable than young drivers. The least amount of disparity in age group variability is evident only with *T-Junction* sign. Also for the older drivers, a definite decrease in CoV is seen for signs with increasing legibility distance. Another observation evident is that the young drivers are more or less uniform in their variability, regardless of sign. The decrease in CoV as a function of longer legibility distances is not seen for young drivers.

Duncan's New Multiple Range Test was performed to determine which specific sign legibility distance means differ from one another. Fifteen pairs of signs were tested. With alpha = 0.05, of the 15 pairs, only two signs pairs were not significantly different with regards to the response variable, legibility distance: *Deer Crossing* and *Slippery When Wet* and *Crossroad* and *T-Junction*. All other 13 pairs were significantly different in mean legibility distance.

Referring to Table 13, the Age X Sign interaction was not significant, implying that although the older group had an overall larger mean legibility distance, there was no particular sign which was much easier or harder for them to read than others, as compared to the younger group.

Subject[Gender,Age] was the error term used to test the between subject effects of gender, age and Gender X Age. Sign X Subject[Gender,Age] was the error term used to test the within subject effects of sign, Gender X Sign, Age X Sign and Gender X Age X Sign.

A Pearson Product Moment correlation was conducted on the mean legibility distance and CoV by sign and by age group. For young drivers, the correlation coefficient was -0.7335, indicating a smaller variability associated with the signs with longer legibility distances. Likewise, for the older drivers, the correlation coefficient was even higher: -0.9032. This implies that the negative strength of the relationship between variation and legibility distance is even stronger for the older drivers. For the signs for which they have particularly short legibility distances (e.g., *Bicycle Crossing*), the coefficient of variation is very high (for *Bicycle Crossing*, the CoV for older drivers was 62.8%).

Analysis of Trials for Learning

In order to see whether the six replications were resulting in subjects' learning, a Duncan's New Multiple Range Test was performed on the six trial means by sign. The data was analyzed by taking the means for trials one through six for four subgroups of data and then applying Duncan's test for differences between the adjacent paired means. The four subgroups were by sex and age, or older females, older males, young males and young females. Although there was a weak positive slope to distance as a function of trials, the Duncan's test showed the adjacent pairs compared to be non-significant. Table 16 shows by sign, the ordered legibility distances by trial. Table 16 does show that for four of the six signs, the shortest legibility distance occurs for trials one or two. The longest legibility distance is obtained on either trial five or six for five of the six signs.

TABLE 16

Rank order legibility distance by trial by sign, field data only

					
6	6	6	6	5	6
5	5	5	4	6	5
4	4	4	5	4	2
3	3	3	2	3	3
2	1	2	1	2	4
1	2	1	3	1	1

LABORATORY STUDY

The same ANOVA described above was used to test the laboratory data. JMP® Version 3.0.2, once again ran the equivalent of PROC MIX on the laboratory legibility distances. The same model statement for the field data was used for the laboratory data.. Table 17 is the source table for the analysis performed.

TABLE 17

ANOVA source table for laboratory study data only

Source	SS	MS	df	F Ratio	Prob>F
Gender	1173735.7	1173735.7	1	0.47	0.5001
Age	26170305	26170305	1	10.51	0.0041*
Gender*Age	4585048.8	4585048.8	1	1.84	0.1898
Sign	24847072	4969414	5	54.17	0.0000*
Gender*Sign	559380.5	111876	5	1.22	0.3055
Age*Sign	153168	30633.6	5	0.33	0.8913
Gender*Age*Sign	450870	90174	5	0.98	0.4320
Subj[Gender,Age]	49778975.0	2488948.8	20	27.13	0.0000
Sign*Subj[Gender,Age]	9172959.6	91729.6	100	5.04	0.0000

*** denotes significance at alpha=0.05**

As in the field study, Table 17 reveals significant main effects only for the independent variables of age and individual warning symbol sign for the laboratory study results. The main effect for the between subject variable of gender was not significant in the laboratory portion of this research either. Males and females did not differ in their legibility distance responses to any of the six signs. As for the field study, females saw the signs slightly farther away than males. The average legibility distance for females was 325.4 meters (1067.9 feet), and was 311.5 meters (1022.0 feet) for males.

Figure 25 is a plot of the age main effect result. As in Figure 22, Figure 25 plots means legibility distance (in meters). Once again, the young participants had an overall longer legibility distance for all signs. The young subjects had a mean legibility distance of 372.2 meters (1221.3 feet); older participants had a mean legibility distance of 263.6 meters (864.7 feet). The legibility distance for older subjects was approximately 71% that of younger drivers.

Just as in the field study, the main effect of individual warning symbol sign was significant. The six signs did differ in their mean legibility distances across all subjects. Figure 26 plots the mean legibility distance for the six signs.

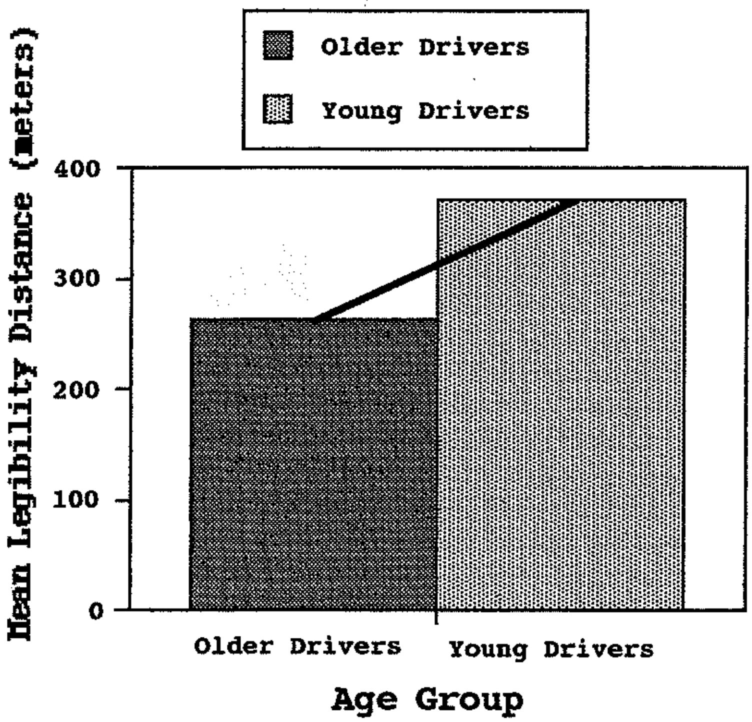

Figure 25. *Main effect plot of age for laboratory study*

Table 18 below summarizes the mean legibility distance for each sign and the standard deviation. Additionally, the standard error of the mean is contained in the table.

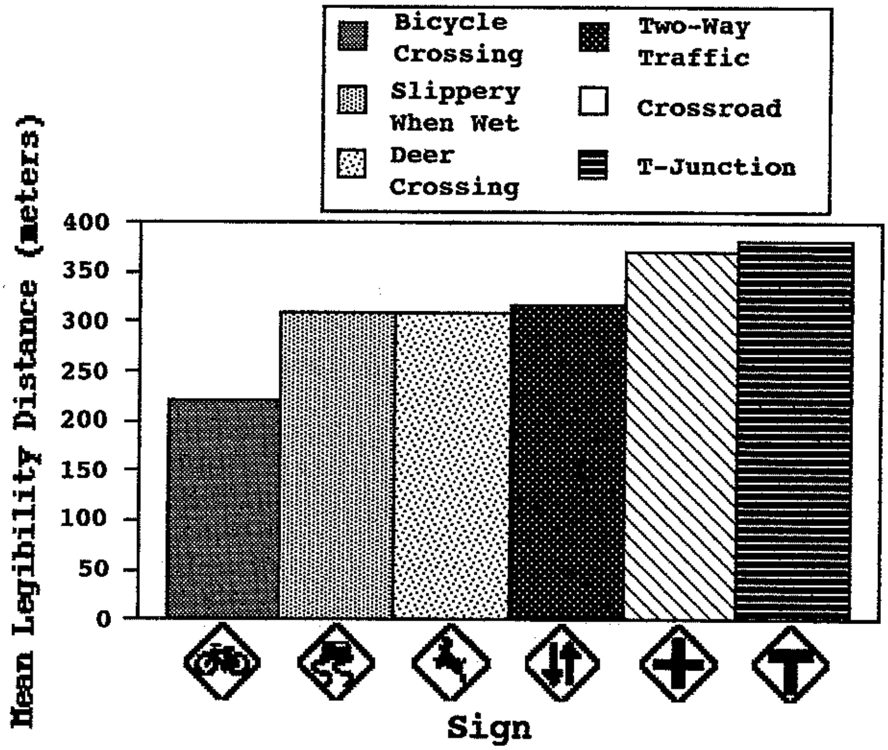

Figure 26. *Main effect plot of individual warning sign, laboratory data*

A comparison for each pair of signs was made using Duncan's Multiple Range Test for the variable legibility distance. Of the 15 pairs of signs tested, only four were not significantly different from one another in legibility distance. The four pairs are: *Crossroad* and *T-Junction*, *2-Way Traffic* and *Deer Crossing*, *2-Way Traffic* and *Slippery When Wet*, and *Deer Crossing* and *Slippery When Wet*. All other pairs of signs differed significantly in legibility distance.

As in the field study, the sign which had the longest mean legibility distance was *T-Junction*. Likewise, as in the field study, *Bicycle Crossing* was the warning sign which had the shortest mean legibility distance across all subjects. Table 19, CoV by symbol sign, is derived using the mean legibility distance data and the standard deviation.

TABLE 18

Mean, standard deviation and standard error of mean for individual sign legibility distance, laboratory study only

Warning Sign	Mean Legibility Distance *	Standard Deviation *	Standard Error of Mean *
Bicycle Crossing (W11-1)	220.45 m (723.26)	94.15 m (308.884)	7.85 m (25.740)
Slippery When Wet (W8-5)	308.27 m (1011.39)	111.66 m (366.348)	9.31 m (30.529)
Deer Crossing (W11-3)	308.82 m (1013.19)	110.91 m (363.860)	9.24 m (30.322)
Two-Way Traffic (W6-3)	317.57 m (1041.88)	106.52 m (349.487)	8.88 m (29.124)
Crossroad (W2-1)	371.41 m (1218.54)	112.30 m (368.421)	9.36 m (30.702)
T-Junction (W2-4)	381.00 m (1250.00)	107.75 m (353.496)	8.98 m (29.458)

*** Values in parentheses are expressed in feet**

TABLE 19

Coefficient of Variation for six symbol signs, laboratory data only

Warning Sign	Coefficient of Variation (%)
Bicycle Crossing (W11-1)	42.71
Slippery When Wet (W8-5)	36.22
Deer Crossing (W11-3)	35.91
Two-Way Traffic (W6-3)	33.54
Crossroad (W2-1)	30.24
T-Junction (W2-4)	28.28

Similar to the field study results, the CoVs for laboratory legibility distances for all six warning signs are also very large, ranging from over 28% to almost 43%. Exactly as in the field study data, the sign with the largest CoV was *Bicycle Crossing*. The sign with the smallest variation in legibility distance response was *T-Junction*, another sign which has bold, thick lines, much like the *Crossroad* sign. Recall that the *Crossroad* symbol sign had the smallest CoV in the field study results.

Taking a closer examination of the CoV by sign and by age group, Figure 27 plots the variability as a percentage for both age groups.

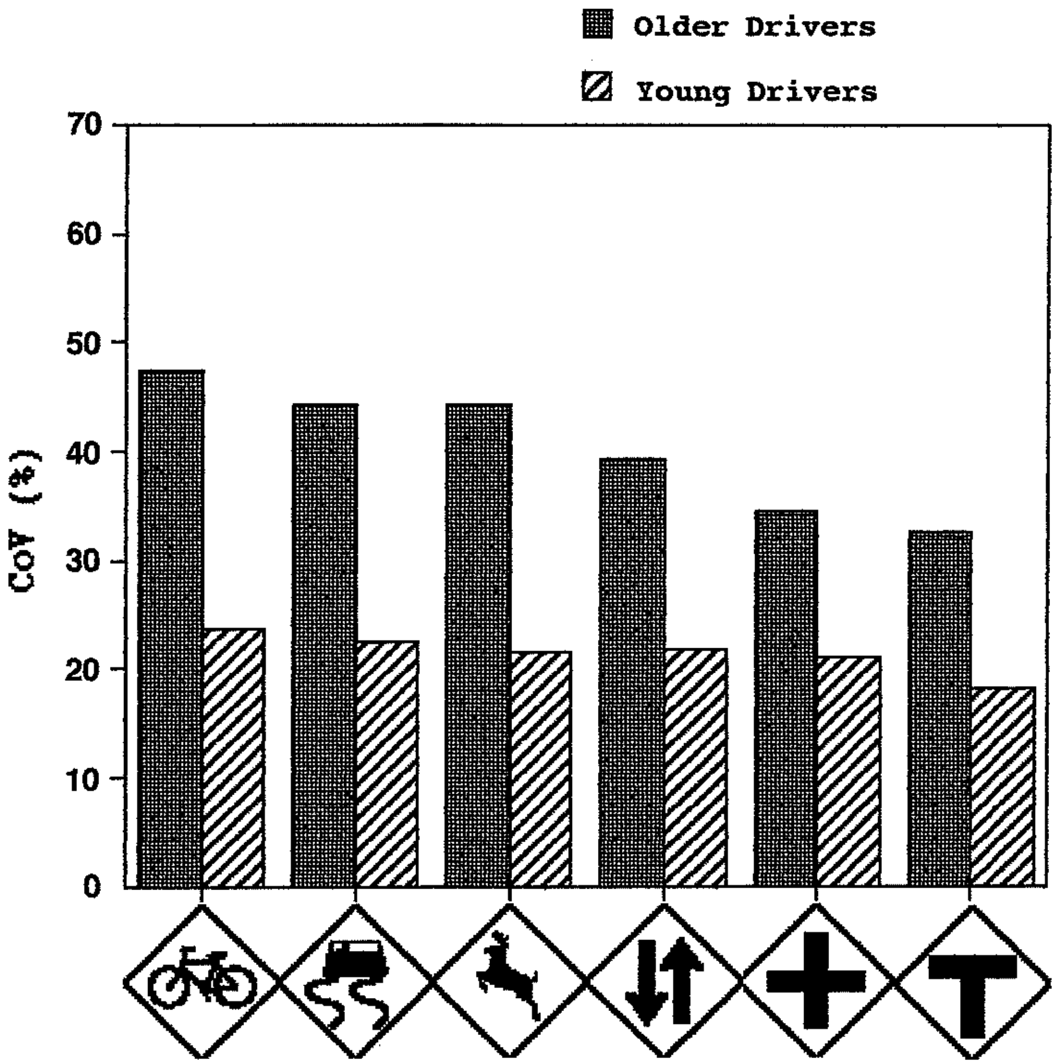

Figure 27. *Plot of CoV by age group, all signs, laboratory data*

As noted for the field data, for all signs, older drivers were more variable than young drivers. For laboratory data, as seen above in Figure 27, the CoV for older drivers is more than twice that of the young group for three signs: *Bicycle Crossing*, *Slippery When Wet* and *Deer Crossing*. The field data found this doubling of CoV for only *Bicycle Crossing* and *Slippery When Wet* (see Figure 24). Exactly as for the field data, older drivers' CoV decreases as the legibility distance for the sign increases. Again, the young drivers are nearly invariant in their CoV percentages, regardless of sign.

A Spearman rank correlation coefficient was calculated for the symbol signs rank (by legibility distance) between the field and laboratory data results. The ranks were exactly the same between the two portions, therefore, the r_S was equal to 1.00.

Table 20 catalogs by age and individual sign, the mean legibility distances, plus standard deviations. Also included in the table is the CoV for each sign's legibility distance by age group. It is these CoV for age groups by sign which are plotted above in Figure 27.

TABLE 20

Mean legibility distance, standard deviation and CoV for signs by age groups - laboratory data

Warning Sign	Mean Legibility Distance - Young*	Mean Legibility Distance- Older*	Standard Deviation Young*	Standard Deviation Older*	CoV Young %	CoV Older %
Bicycle Crossing (W11-1)	282.2 m (925.8)	158.7 m (520.7)	66.7 m (218.7)	75.3 m (247.2)	23.6	47.5
Deer Crossing (W11-3)	359.4 m (1179.0)	258.3 (847.4)	80.3 m (263.5)	114.6 m (376.1)	22.3	44.4
Slippery When Wet (W8-5)	364.8 m (1196.7)	251.8 (826.1)	78.4 m (257.3)	111.6 m (366.7)	21.5	44.4
Two-Way Traffic (W6-3)	369.8 m (1213.2)	265.4 m (870.6)	80.7 m (264.8)	104.0 m (341.1)	21.8	39.2
Crossroad (W2-1)	422.9 m (1387.4)	319.9 m (1049.7)	88.7 m (291.1)	110.3 m (361.8)	21.0	34.5
T-Junct (W2-4)	434.7 m (1426.0)	327.4 (1074.0)	78.8 m (258.7)	106.4 m (349.2)	18.1	32.5

*** Values in parentheses are expressed in feet**

For the sign *Bicycle Crossing*, the young participant could see and correctly identify the sign by a distance almost twice that of the older driver (1.78 times the distance). The for other five signs, younger subjects could see *Crossroad* at a distance of 1.32 times that of older drivers, *T-Junction* at 1.33 times, both *Two-Way Traffic* and *Deer Crossing* were seen at 1.39 times farther, and *Slippery When Wet* was seen at a distance of 1.45 times farther than older subjects. The young subjects, as a group, are less variable in their responses to the six signs than the older participants, by virtue of the lower values for CoV. In the case of three of the signs, (*Bicycle Crossing, Deer Crossing,* and *Slippery When Wet*) the CoV for the older driver legibility distances was over twice that of the young driver.

As described above, the two terms Subject[Gender,Age] and Sign*Subject[Gender,Age] served as the error terms to test the between and within subject variables, respectively.

Just as described above for field data, the laboratory mean legibility distances and CoV by age group for each sign were correlated. The Pearson Product Moment correlation coefficient for young drivers was -0.8760, while the same correlation coefficient for older drivers was -0.9044. Like the discussion for the field data above, the variances are inversely related to the legibility distances by sign by age group. Both correlations are negative and large, showing a strong negative association of variance with mean legibility distance. For both age groups, but especially for the older drivers, the signs with the shortest legibility distance have the most variability associated. On the other hand, signs with long legibility distances have smaller variability associated with the responses.

Analysis of Trials for Learning

As described above for the field data, the six replications were analyzed using Duncan's New Multiple Range Test to investigate a possible learning effect occurring. The data was analyzed by the same method as used for field data for each sign. Means for trials one through six for four subgroups of data were computed and Duncan's test for differences between the adjacent paired means was performed. The four subgroups were by sex and age: older females, older males, young males and young females. Duncan's test showed the adjacent pairs compared to be non-significant. Table 21 shows by sign, the ordered legibility distances by trial for the laboratory data. Table 21 does show that for only two of the six signs, the shortest legibility distance occurs for trials one or two. The longest legibility distance is obtained on either trial five or six for three of the six signs.

When comparing Tables 16 and 21, examining the field and laboratory rank order legibility distances by sign, it is seen that the rank order of legibility distances is much different for laboratory data than for field data. In the field, although Duncan's test showed there were no significant adjacent pairs and hence, no learning ongoing, the rank order of trials were much like one would expect if learning was going on across the six trials. In other words, the order was ascending over trials from shortest to longest legibility distances. An examination of Table 21 shows the laboratory data is not exhibiting the same trend over trials as seen in the field data (Table 16).

TABLE 21

Rank order legibility distance by trial by sign, laboratory data only

					
6	1	6	6	5	5
5	5	3	4	6	6
3	4	4	5	4	1
4	6	5	1	2	3
1	2	2	2	1	2
2	3	1	3	3	4

ALL DATA - FIELD AND LABORATORY

Above, individual data sets were analyzed separately, one set of analyses for field and another for the laboratory study. Data sets from both studies were merged together and another PROC MIXED analysis was performed, this time adding the main effect variable of condition (field or laboratory) and all interactions with condition. The ratio of the model total mean square error for the laboratory and field data separately was 1.088. The ratio of the variance the two conditions' data was less than 3.84, therefore, the data sets could be combined. Table 22 is the source table from that data analysis run.

TABLE 22

ANOVA source table for all data, field and laboratory combined

Source	S S	MS	df	F Ratio	Prob>F
Gender	1354034	1354034	1	0.28	0.6007
Age	52205620	52205620	1	10.91	0.0036*
Gender*Age	9524410	9524410	1	1.99	0.1737
Sign	46346183	9269237	5	91.85	0.0000*
Gender*Sign	870436	174087	5	1.73	0.1357
Age*Sign	417376	83475.2	5	0.83	0.5333
Gender*Age*Sign	501646	100329	5	0.99	0.4253
Condition	4426884	4426884	1	21.05	0.0002*
Gender*Condition	135804	135804	1	0.65	0.4311
Age*Condition	87.15	87.15	1	0.00	0.9840
Sign*Condition	966811	193362	5	3.53	0.0055*
Gender*Age*Cond	3357.9	3357.9	1	0.02	0.9007
Age*Sign*Cond	427136	87427.2	5	1.56	0.1779
Gender*Sign*Condition	194379	38875.9	5	0.71	0.6168
Gender*Age*Sign*Condition	209255	41851	5	0.77	0.5771
Subj[Gender,Age]	95719258	4785963	20	18.66	0.0000
Sign*Subj[Gender,Age]	10091611	100916	100	1.84	0.0012
Condition*Subj[Gender,Age]	4206772	210339	20	3.85	0.0000
Sign*Cond*Subj[Gender,Age]	5470504	54705	100	2.50	0.0000

Scrutinizing Table 22, it is not surprising to see that the main effects of age and sign are significant. These two main effects were both significant when the data results were analyzed separately by field and laboratory sessions. Analyzing all 1728 observations (24 subjects x 6 signs x 6 replications per sign x 2 conditions), the main effect of condition (field or laboratory) was highly significant ($p<0.0002$). Furthermore, the two-way interaction of Sign X Condition was also significant ($p<0.0055$).

The average legibility distance in the field for all subjects and all signs was 287.4 meters (942.8 feet), while the same distance for the laboratory portion was 317.9 meters (1043.0 feet). This difference in average legibility distance for the two conditions was significant. Figure 28 shows the plot of mean legibility distance for the Sign X Condition interaction.

Table 23 catalogs the mean legibility distance in meters (and feet) for all six signs, by condition, along with the difference, expressed as a percentage. From Figure 28 above, one can see that in all cases, all signs were legible from a farther distance in the laboratory than in the field session. By way of summarizing the data found in Tables 14 and 18, the mean legibility distances are repeated in Table 23.

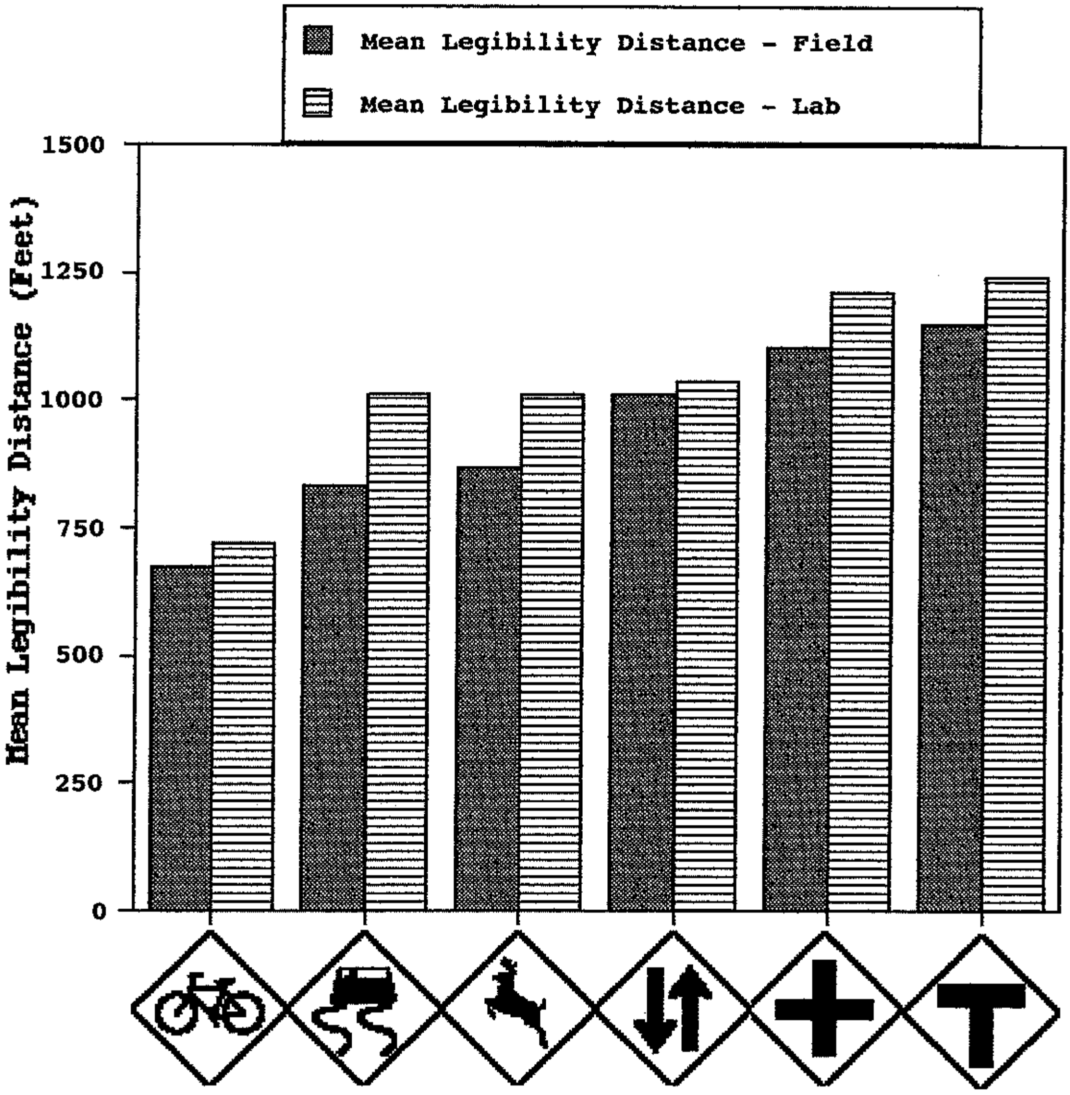

Figure 28. *Sign*Condition interaction plot*

TABLE 23

Mean legibility distance and delta (percentage) by condition for all signs

Warning Sign	**Mean Legibility Distance-Field***	**Mean Legibility Distance-Lab***	**Percentage Difference Between Lab and Field**
Bicycle Crossing (W11-1)	205.0 (672.6)	220.5 (723.3)	+ 7.5%
Slippery When Wet (W8-5)	254.4 (834.5)	308.3 (1011.4)	+21.2%
Deer Crossing (W11-3)	270.2 (886.5)	308.8 (1013.2)	+14.3%
Two-Way Traffic (W6-3)	308.3 (1011.4)	317.6 (1041.9)	+ 3.0%
Crossroad (W2-1)	336.2 (1102.9)	371.4 (1218.5)	+10.5%
T-Junction (W2-4)	350.3 (1149.1)	381.1 (1250.0)	+ 8.8%
		Average Difference	**+ 10.9%**

* **values in parentheses are in feet**

Slippery When Wet has the largest difference in mean legibility distances (+21.2% farther) between conditions, however, there is little difference between field and laboratory sessions for *Two-Way Traffic*. As shown in Table 23, the average overall difference between field and laboratory mean legibility distances is +10.9%, indicating that for all signs, they are seen farther in the laboratory than the field, by an average of less than 11%.

Overall ranges of responses for legibility distance by sign and condition , regardless of age group are summarized in Table 24. This range reporting begins to hint at the large within subject variability present for both studies.

TABLE 24

Minimum and maximum legibility response by sign and condition

Warning Sign	Minimum Field Response*	Maximum Field Response*	Minimum Lab Response*	Maximum Lab Response*
Slippery When Wet (W8-5)	42.6 (139)	470.0 (1542)	73.2 (240)	533.4 (1750)
Bicycle (W11-1)	43.6 (143)	516.9 (1696)	67.1 (220)	466.3 (1530)
Deer (W11-3)	45.1 (148)	576.1 (1860)	67.1 (220)	548.6 (1800)
Crossroad (W2-1)	80.8 (265)	536.5 (1760)	76.2 (250)	548.6 (1800)
2-Way (W6-3)	104.5 (343)	530.4 (1740)	73.2 (240)	536.5 (1760)
T-Junction (W2-4)	113.4 (372)	579.1 (1900)	57.9 (190)	542.5 (1780)

*** values in parentheses are in feet**

Figure 29 plots the data appearing in Table 24 for the field condition. Figure 30 is a plot of the same information, but for the laboratory condition. These Figures show the mean plotted with the minimum and maximum responses (in feet) by individual sign.

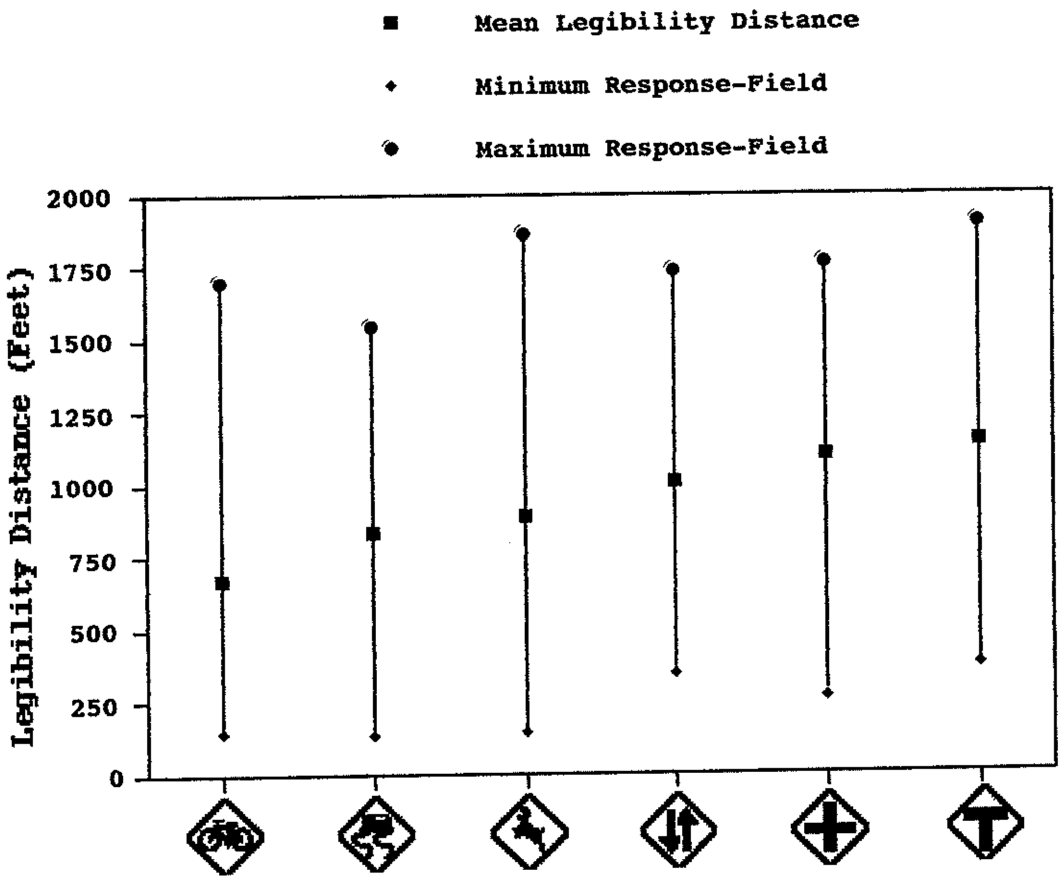

Figure 29. *Mean, minimum and maximum responses by sign, field condition only.*

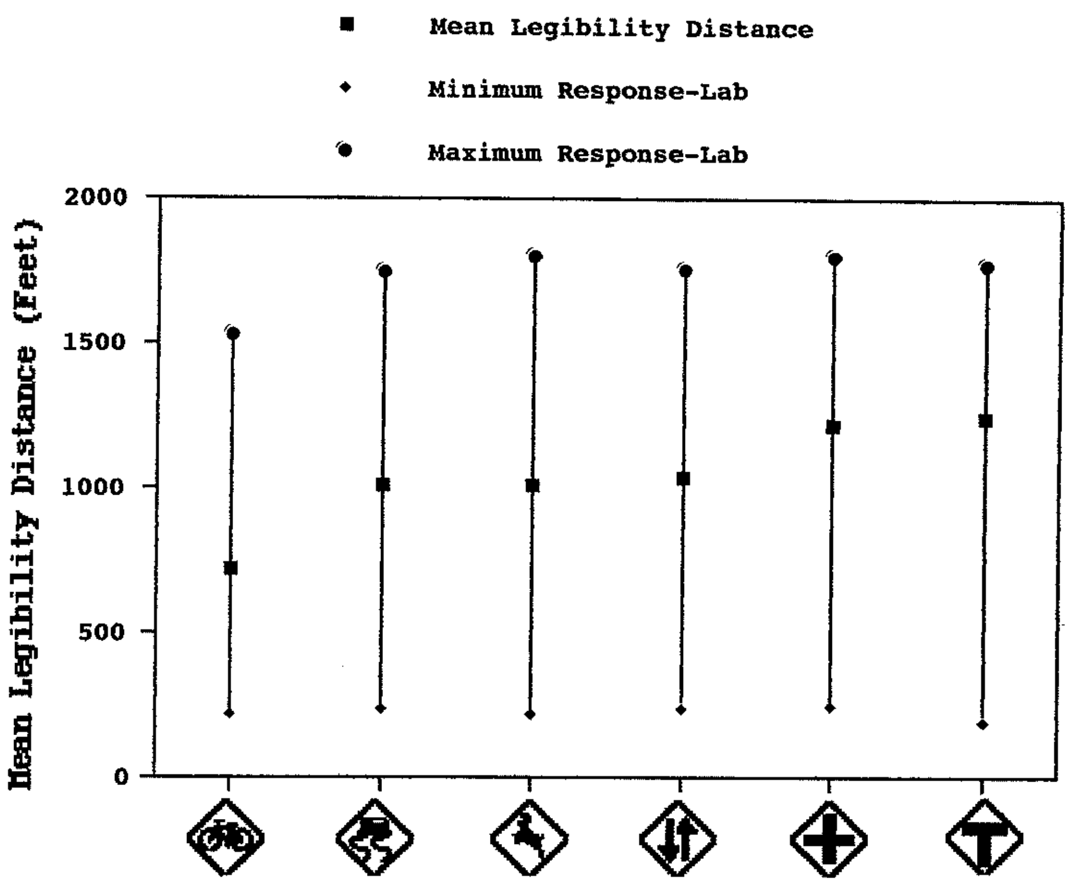

Figure 30. *Mean, minimum and maximum responses by sign, laboratory condition only.*

Tables 25 and 26 show the mean, minimum and maximum response - legibility distance - as a function of age. Table 25 has the data for young drivers, while Table 26 summarizes the information for older drivers.

TABLE 25

Minimum and maximum legibility response by sign and condition, young drivers only

Warning Sign	Minimum Field Response*	Maximum Field Response*	Minimum Lab Response*	Maximum Lab Response*
Bicycle Crossing	106.7 (350)	469.7 (1541)	121.9 (400)	466.3 (1530)
Deer Crossing	157.3 (516)	566.9 (1860)	164.6 (540)	536.5 (1760)
Slippery When Wet	153.3 (503)	466.0 (1529)	216.4 (710)	533.4 (1750)
Two-Way Traffic	184.4 (605)	492.3 (1615)	201.2 (660)	536.5 (1760)
Cross-road	203.3 (667)	509.6 (1672)	231.7 (760)	548.6 (1800)
T-Junction	194.5 (638)	579.1 (1900)	249.9 (820)	539.5 (1770)

*** values in parentheses are in feet**

TABLE 26

Minimum and maximum legibility response by sign and condition, older drivers only

Warning Sign	Minimum Field Response*	Maximum Field Response*	Minimum Lab Response*	Maximum Lab Response*
Bicycle Crossing	43.6 (143)	516.9 (1696)	67.1 (220)	432.8 (1420)
Deer Crossing	45.1 (148)	467.9 (1535)	67.1 (220)	548.6 (1800)
Slippery When Wet	42.4 (139)	470.0 (1542)	73.2 (240)	506.0 (1660)
Two-Way Traffic	104.6 (343)	530.4 (1740)	73.2 (240)	487.7 (1600)
Cross-road	80.8 (265)	536.5 (1760)	76.2 (250)	542.5 (1780)
T-Junction	113.4 (372)	579.1 (1900)	57.9 (190)	542.5 (1780)

*** values in parentheses are in feet**

Of these minimum legibility distances for both field and laboratory conditions, all subjects were in the older driver category. Interestingly, however, for the maximum field legibility distance, five of the six responses were from a driver in the older age category. Only the maximum legibility distance for *Deer Crossing* was given by a young participant. For the laboratory condition, two warning signs, *Deer Crossing* and *T-Junction* had maximum legibility distance responses from older drivers.

Gender is not a significant variable in this analysis either. Females saw the warning signs at an average legibility distance of 307.3 meters (1008.1 feet); males had a mean legibility distance across all signs and conditions of 298.7 meters (980.1 feet). None of the interactions involving gender were significant.

This section summarized the results from both the field and laboratory studies. However, the issue of within subject variability was so overshadowing that it was felt that an entire section was necessary to convey the results of that analysis.

RESULTS

WITHIN SUBJECT VARIABILITY

FIELD STUDY

As discussed above, there was large within subject variability in the legibility distance data for both the field and laboratory studies. Figures 31 through 42 are plots of the six replications for each subject for each of the six warning symbol signs. These plots of individual legibility distance responses are from the field study only. The laboratory data discussion follows this section.

The figures illustrating the data spread within subject are grouped such that the young drivers are plotted first for each sign, followed by the older drivers' data plot directly below. The mean for each subject is designated by a filled circle. Note the spread of data points by age group and by individual sign.

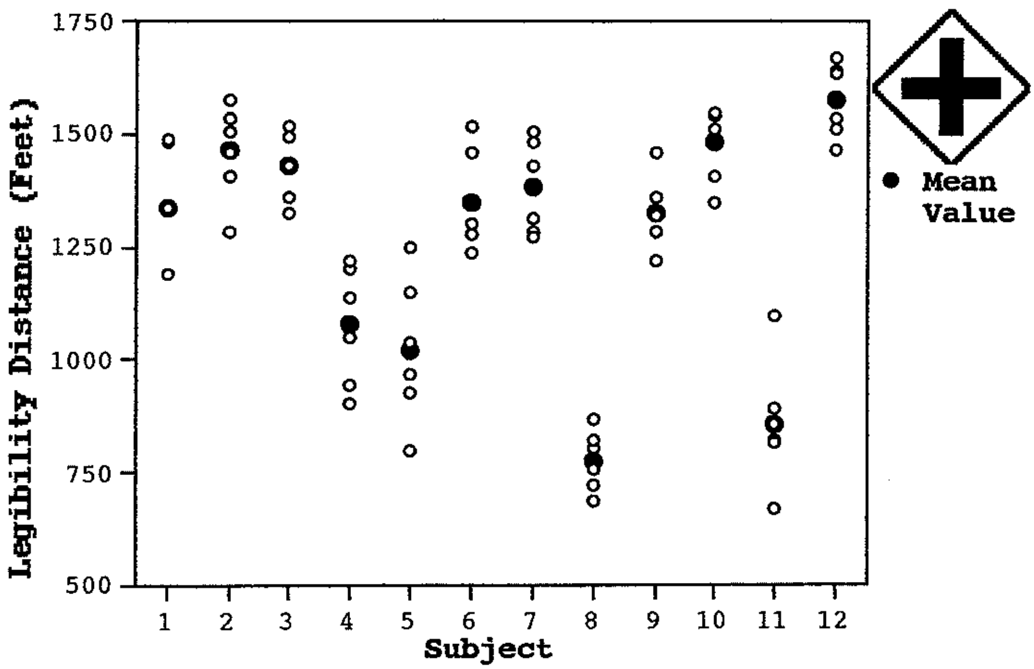

Figure 31. *Field data plot of six trials, with mean. Young drivers and Crossroad (W2-1) warning sign*

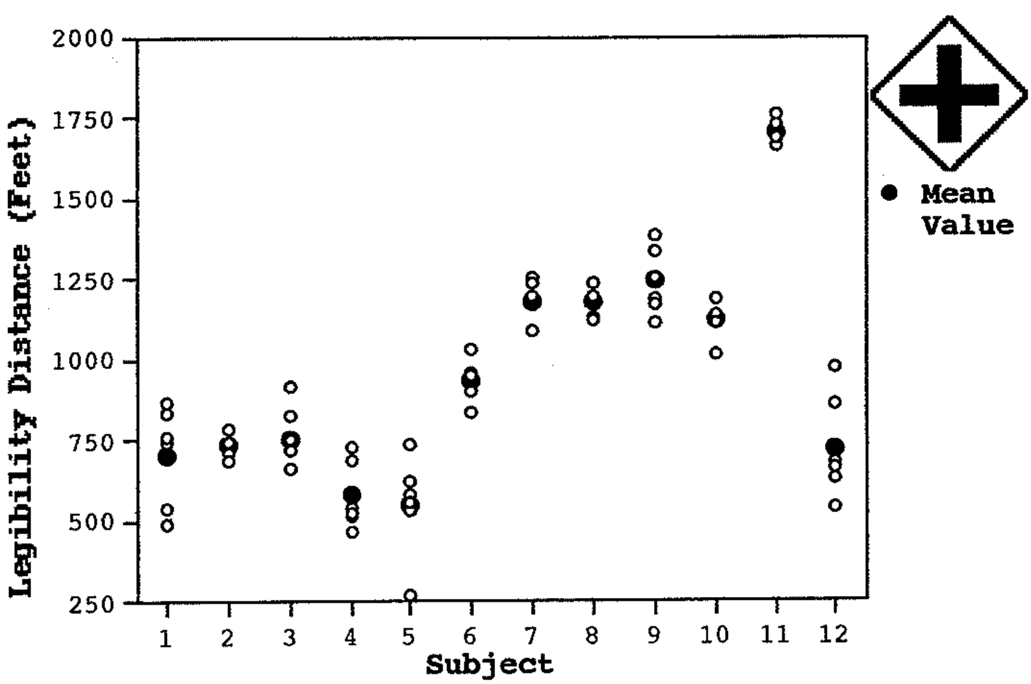

Figure 32. *Field data plot of six trials, with mean. Older drivers and Crossroad (W2-1) warning sign*

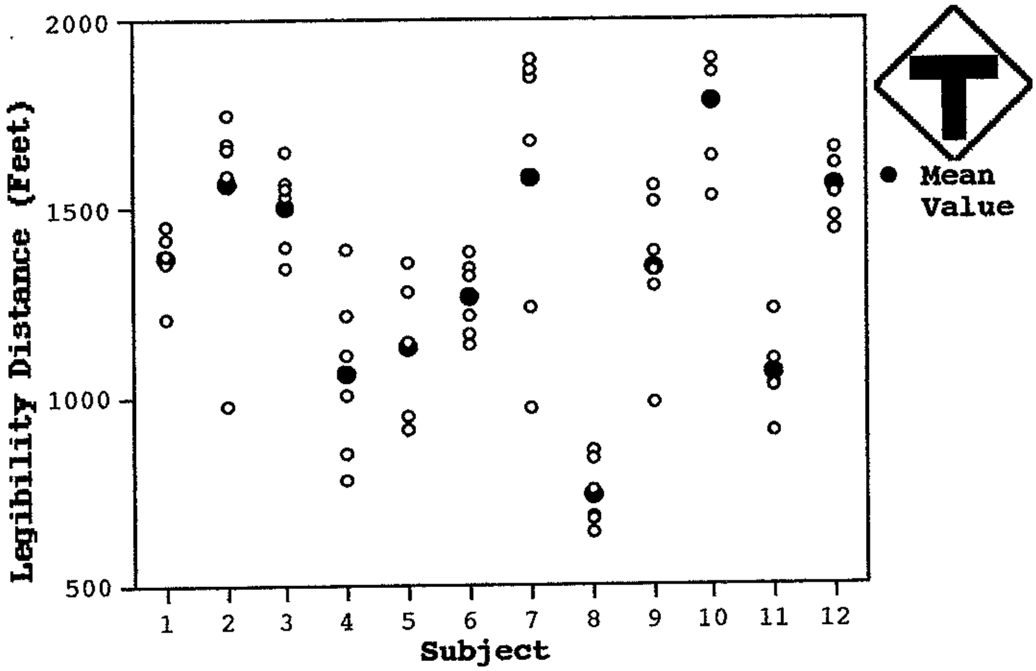

Figure 33. *Field data plot of six trials, with mean. Young drivers and T-Junction (W2-4) warning sign*

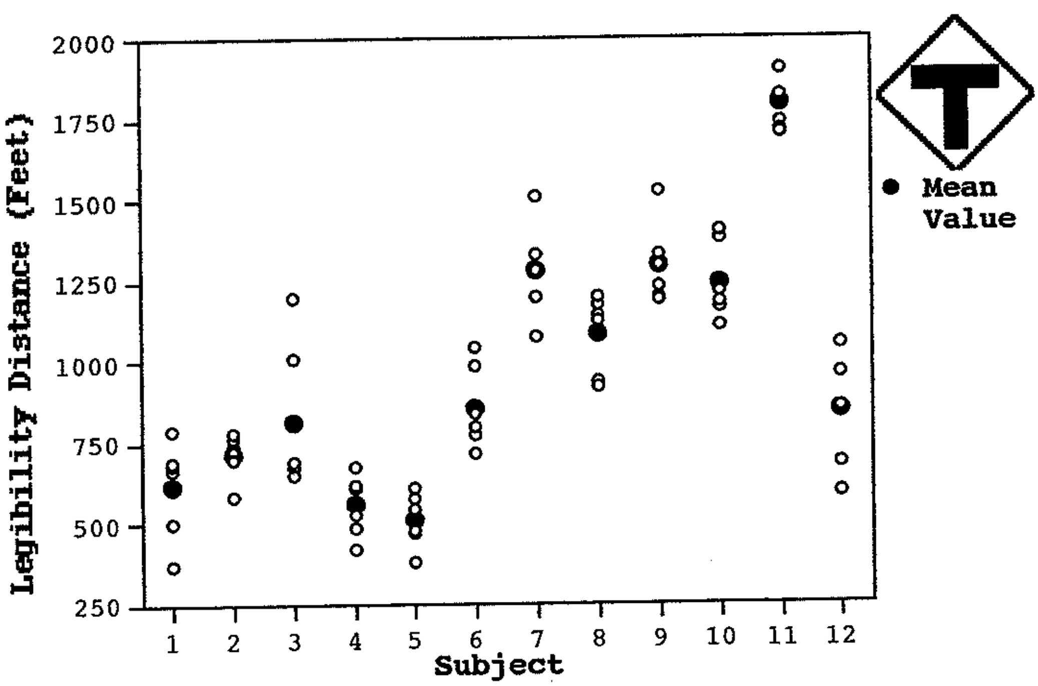

Figure 34. *Field data plot of six trials, with mean. Older drivers and T-Junction (W2-4) warning sign*

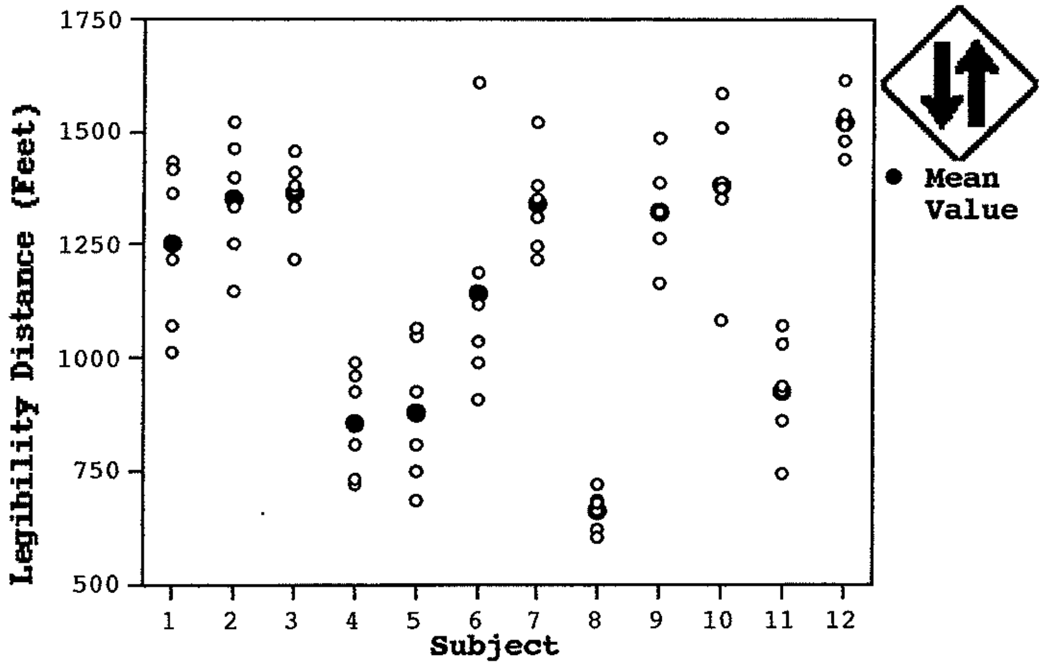

Figure 35. *Field data plot of six trials, with mean. Young drivers and Two-Way Traffic (W6-3) warning sign*

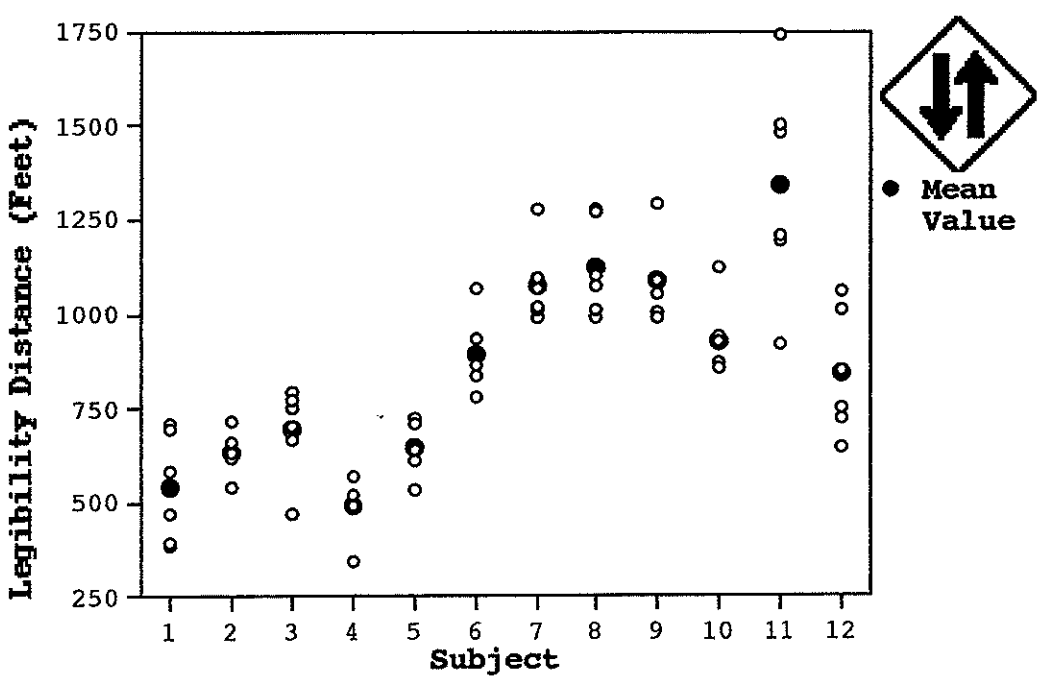

Figure 36 *Field data plot of six trials, with mean. Older drivers and Two-Way Traffic (W6-3) warning sign*

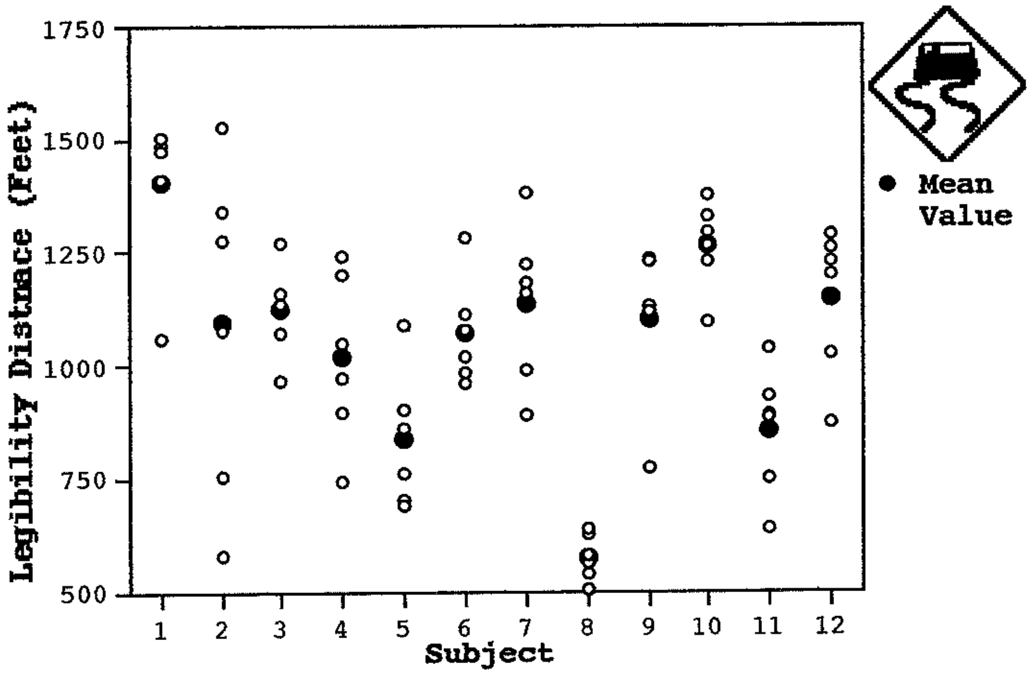

Figure 37. *Field data plot of six trials, with mean. Young drivers and Slippery When Wet (W8-5) warning sign*

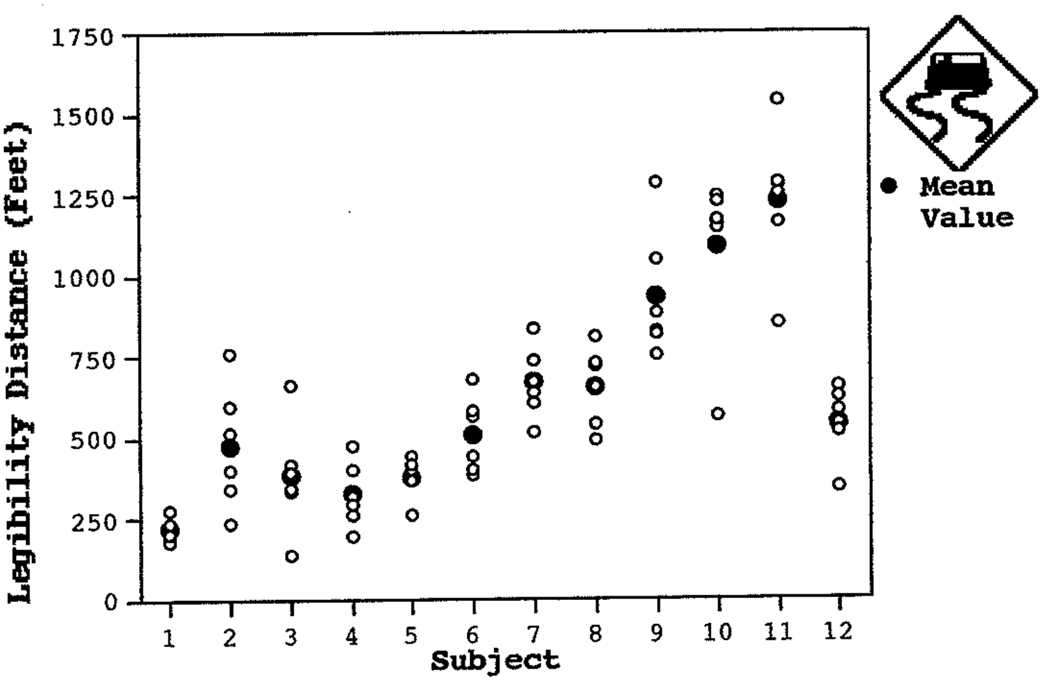

Figure 38. *Field data plot of six trials, with mean. Older drivers and Slippery When Wet (W8-5) warning sign*

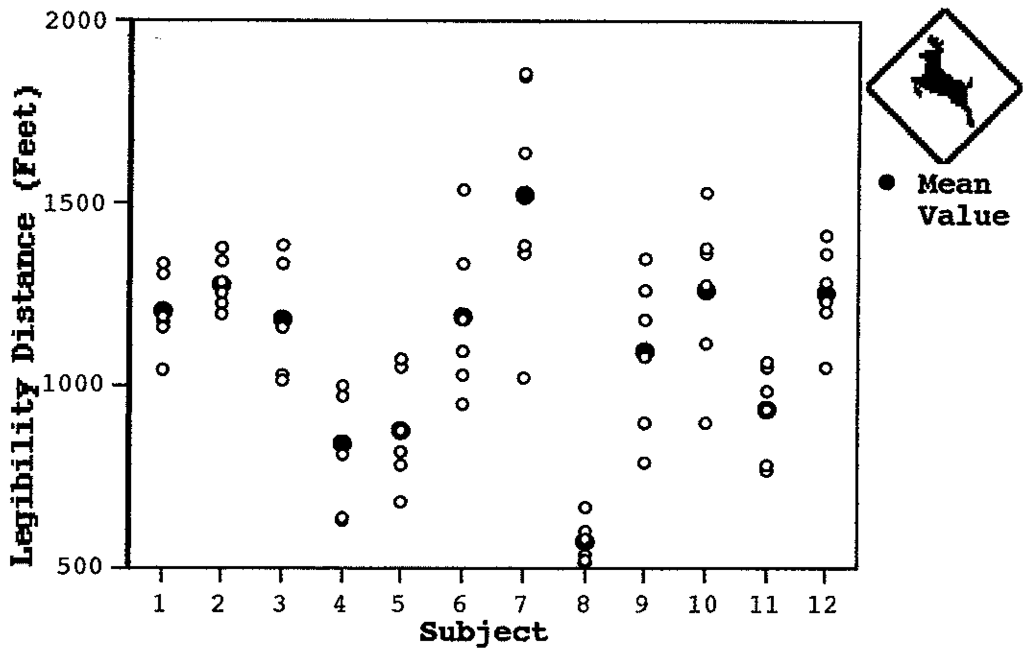

Figure 39. *Field data plot of six trials, with mean. Young drivers and Deer Crossing (W11-3) warning sign*

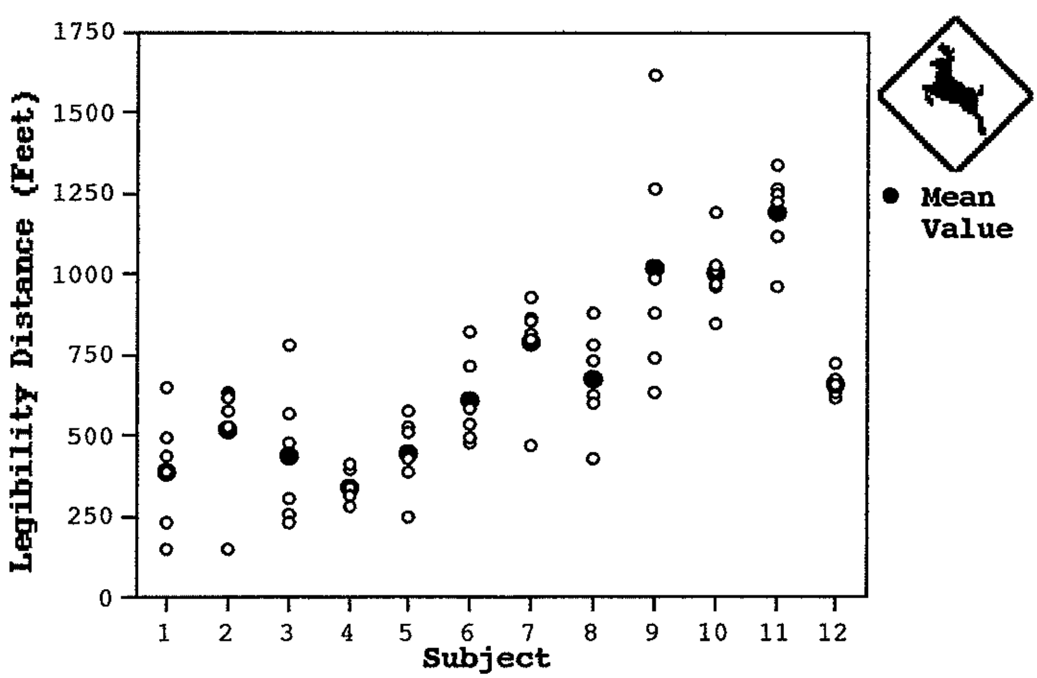

Figure 40. *Field data plot of six trials, with mean. Older drivers and Deer Crossing (W11-3) warning sign*

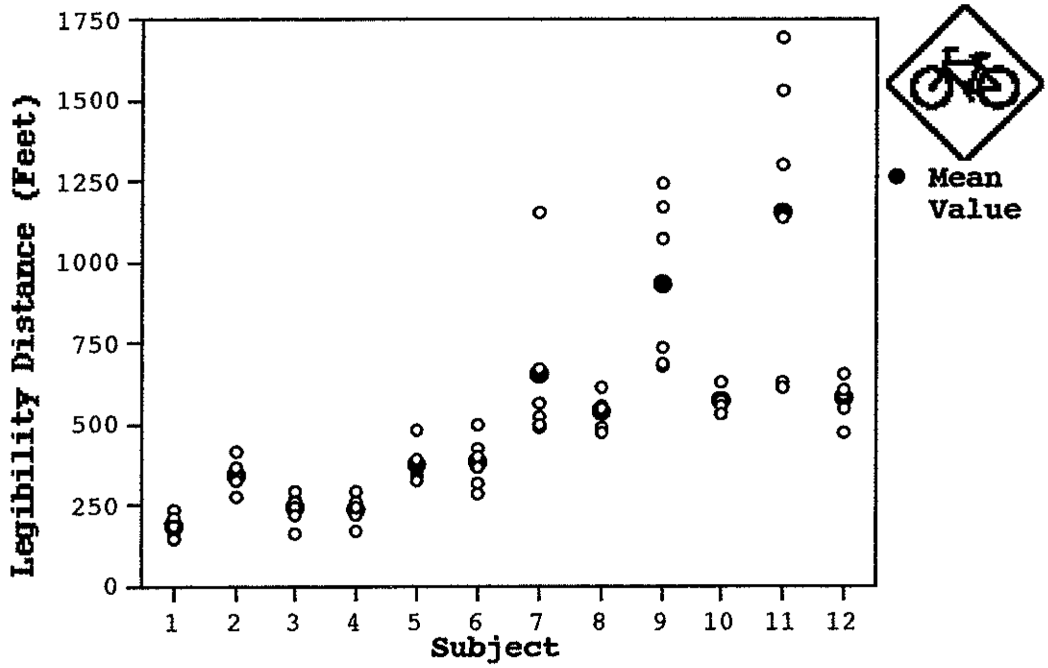

Figure 41. *Field data plot of six trials, with mean. Young drivers and Bicycle Crossing (W11-1) warning sign*

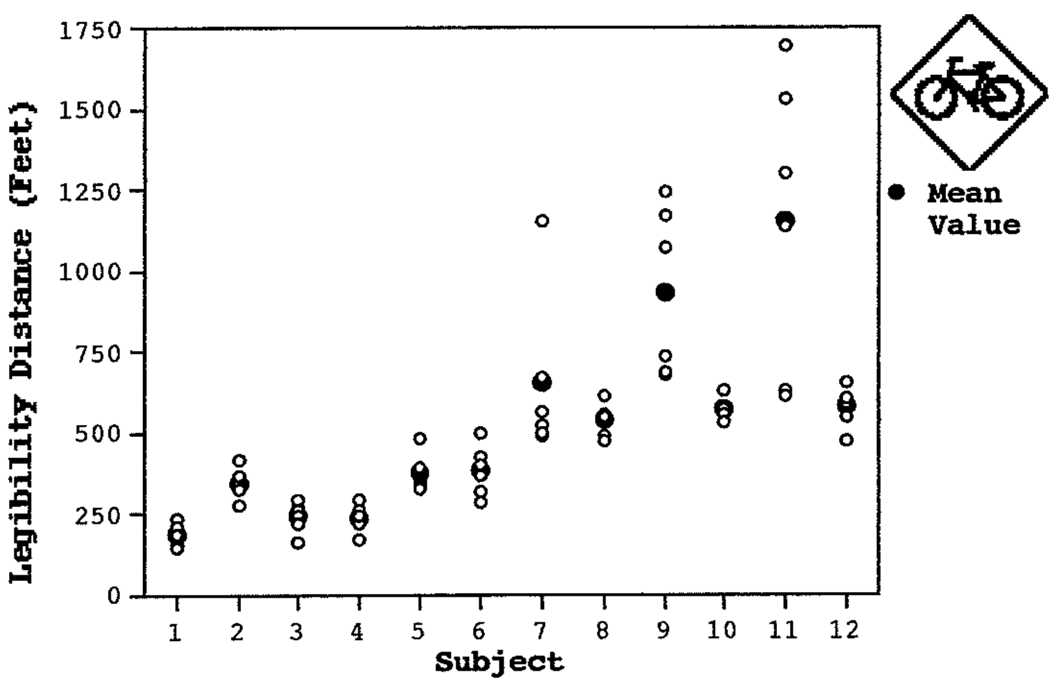

Figure 42. *Field data plot of six trials, with mean. Older drivers and Bicycle Crossing (W11-1) warning sign*

For each subject and sign, a CoV was calculated for each driver over his/her six measures of the response variable, legibility distance. For young drivers, the range of CoVs is shown in Table 27 by individual sign. The ranges for CoV for older drivers are shown in Table 28 for the six warning signs. All data summarized in these tables are for the field study only.

TABLE 27

Ranges of Coefficient of Variation (CoV) by sign, young drivers, field data only

Warning Sign	Minimum CoV	Maximum CoV
Two-Way Traffic (W6-3)	3.9	21.9
Crossroad (W2-1)	5.2	16.3
Deer Crossing (W11-3)	5.4	21.3
T-Junction (W2-4)	5.6	24.6
Bicycle Crossing (W11-1)	6.7	37.0
Slippery When Wet (W8-5)	7.7	33.4

TABLE 28

Ranges of Coefficient of Variation (CoV) by sign, older drivers, field data only

Warning Sign	Minimum CoV	Maximum CoV
Crossroad (W2-1)	2.0	28.6
T-Junction (W2-4)	4.9	28.7
Deer Crossing (W11-3)	5.4	49.2
Bicycle Crossing (W11-1)	5.5	39.4
Two-Way Traffic (W6-3)	8.9	26.7
Slippery When Wet (W8-5)	15.9	44.1

LABORATORY STUDY

The data from the laboratory study also exhibited large within subject variability in the response variable, legibility distance. Figures 43 through 54 are the plots of the six replications per subject, by sign, for both age groups. These figures are analogous to Figures 31 through 42 above, but for the laboratory results. The figures illustrating the data spread within subject are arranged exactly the same as those from the field study. The young drivers' data by individual subject are plotted first for each sign. The second plot is the older drivers' data and falls directly below the young drivers' plot for ease of visual comparison.

The mean for each subject is designated by a filled circle on the data plot. The spread of response variable data points within a subject, by age group and by individual sign is easy to see. These raw data points for legibility distance for each individual subject give a visual sense to the within subject variability.

As outlined above, a CoV was calculated for each subject and sign. Table 29 reports the range of CoVs for young drivers, and the same data is reported for older drivers in Table 30. All data summarized in these tables are for the laboratory study only.

TABLE 29

Ranges of Coefficient of Variation (CoV) by sign, young drivers, laboratory data only

Warning Sign	Minimum CoV	Maximum CoV
Two-Way Traffic (W6-3)	4.3	17.8
Bicycle Crossing (W11-1)	6.8	30.2
Crossroad (W2-1)	3.9	15.8
Deer Crossing (W11-3)	5.1	20.3
Slippery When Wet (W8-5)	3.1	19.0
T-Junction (W2-4)	2.8	15.9

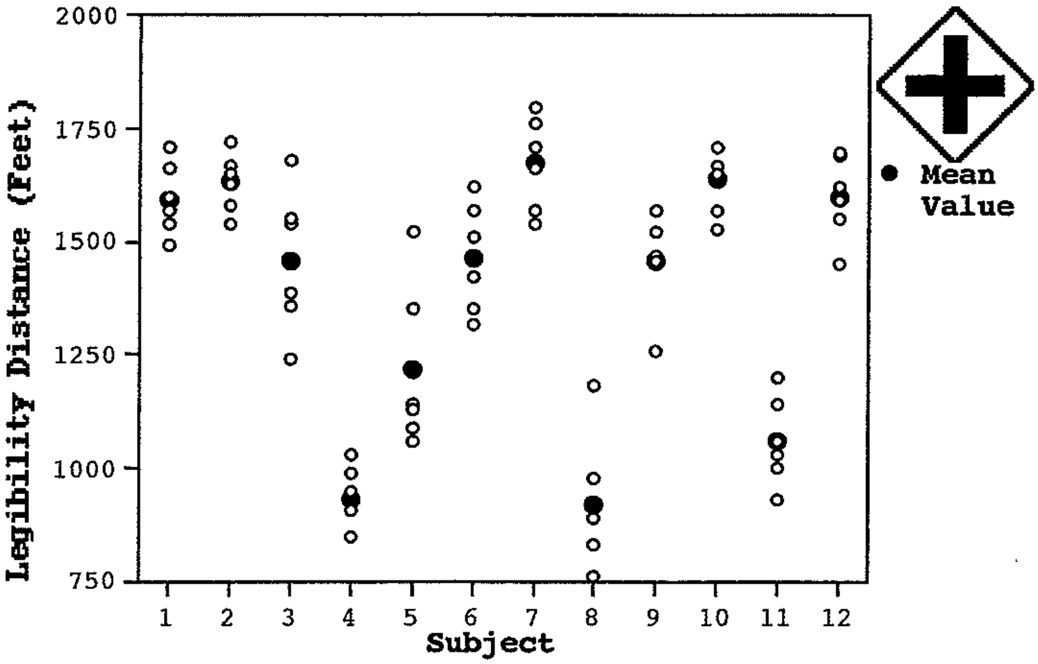

Figure 43. *Laboratory data plot of six trials, with mean. Young drivers and Crossroad (W2-1) warning sign*

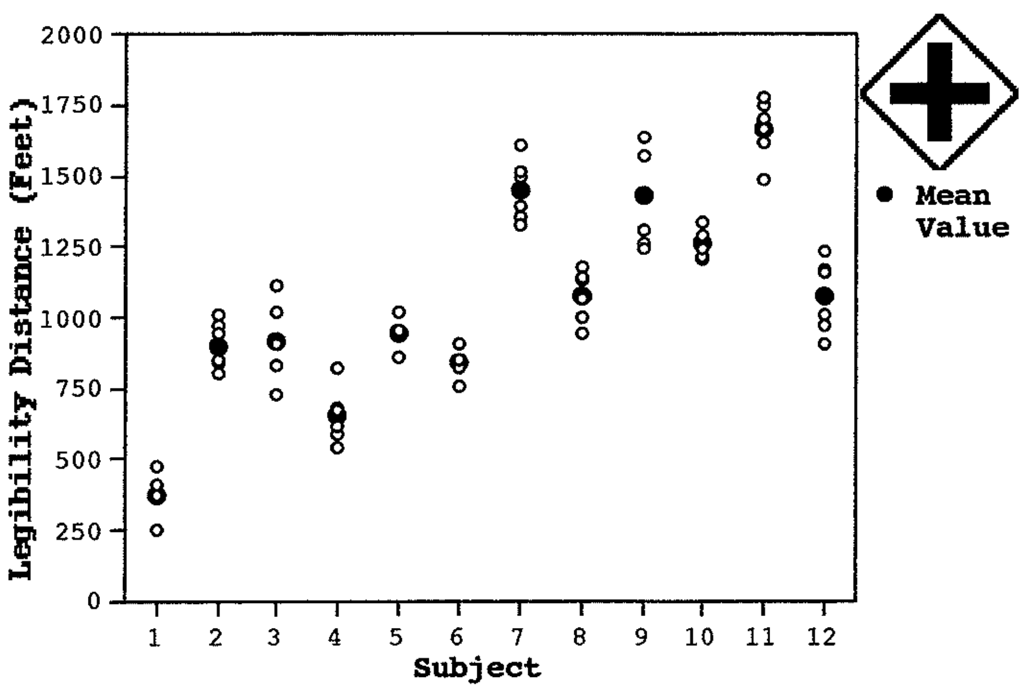

Figure 44. *Laboratory data plot of six trials, with mean. Older drivers and Crossroad (W2-1) warning sign*

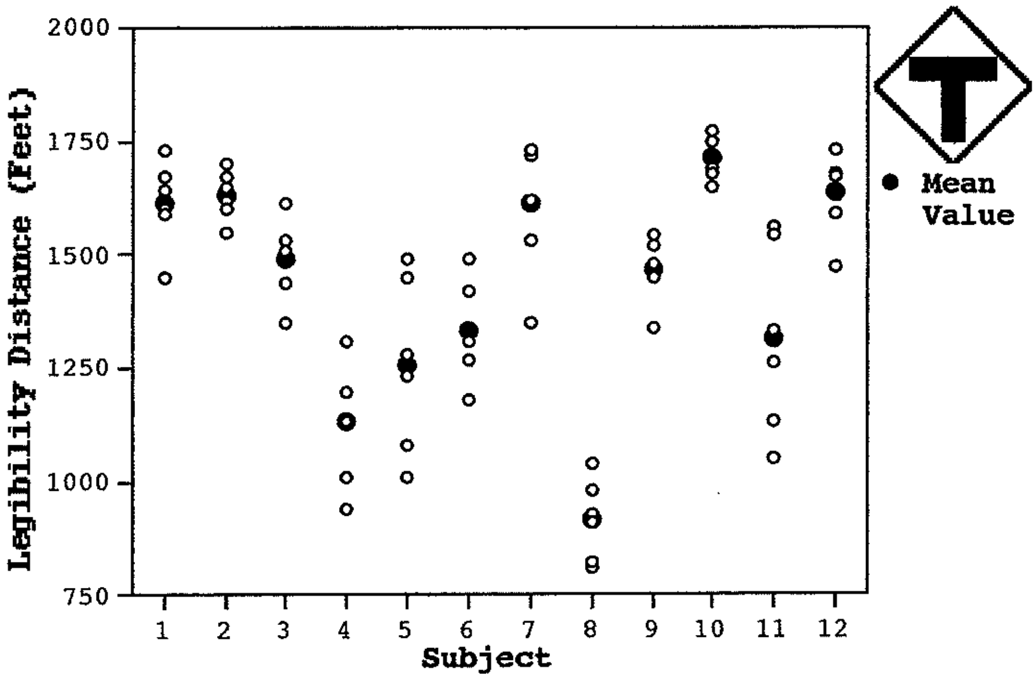

Figure 45. *Laboratory data plot of six trials, with mean. Young drivers and T-Junction (W2-4) warning sign*

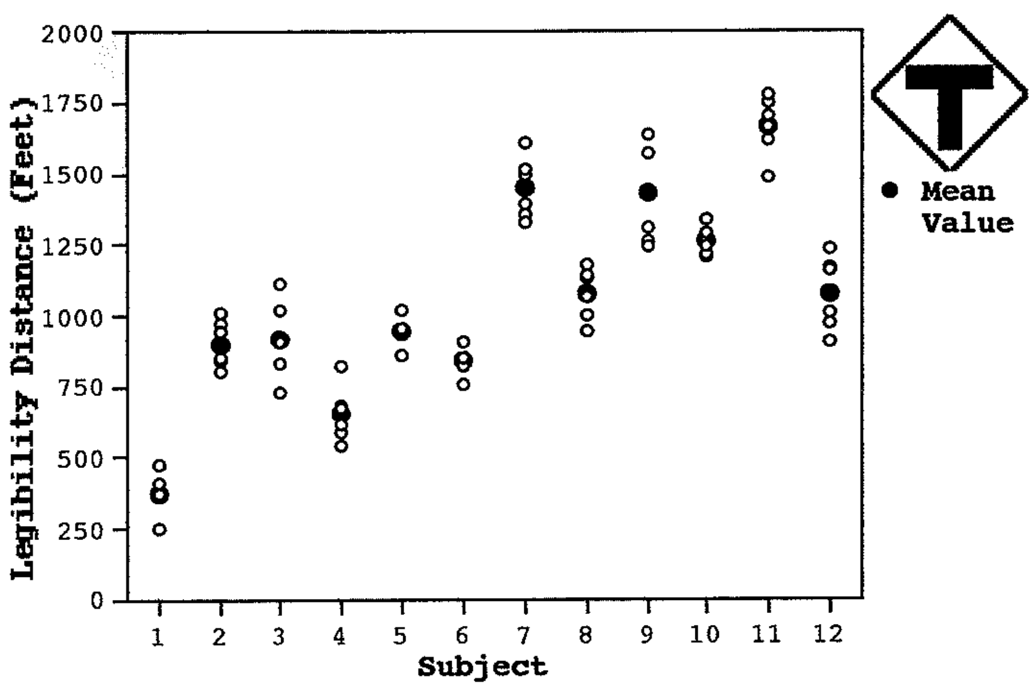

Figure 46. *Laboratory data plot of six trials, with mean. Older drivers and T-Junction (W2-4) warning sign*

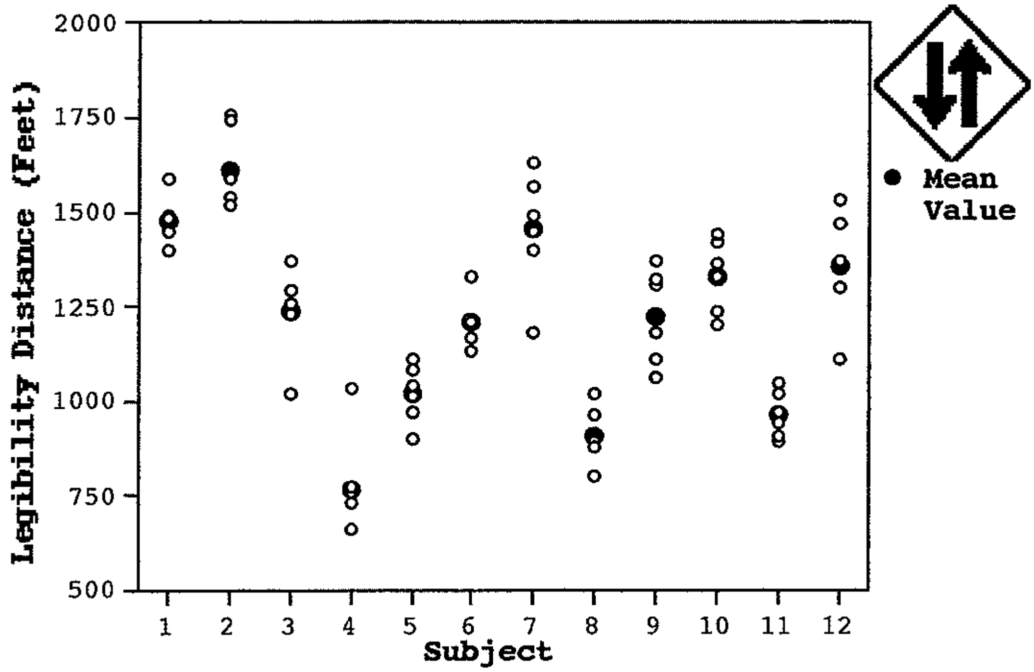

Figure 47. *Laboratory data plot of six trials, with mean. Young drivers and Two-Way Traffic (W6-3) warning sign*

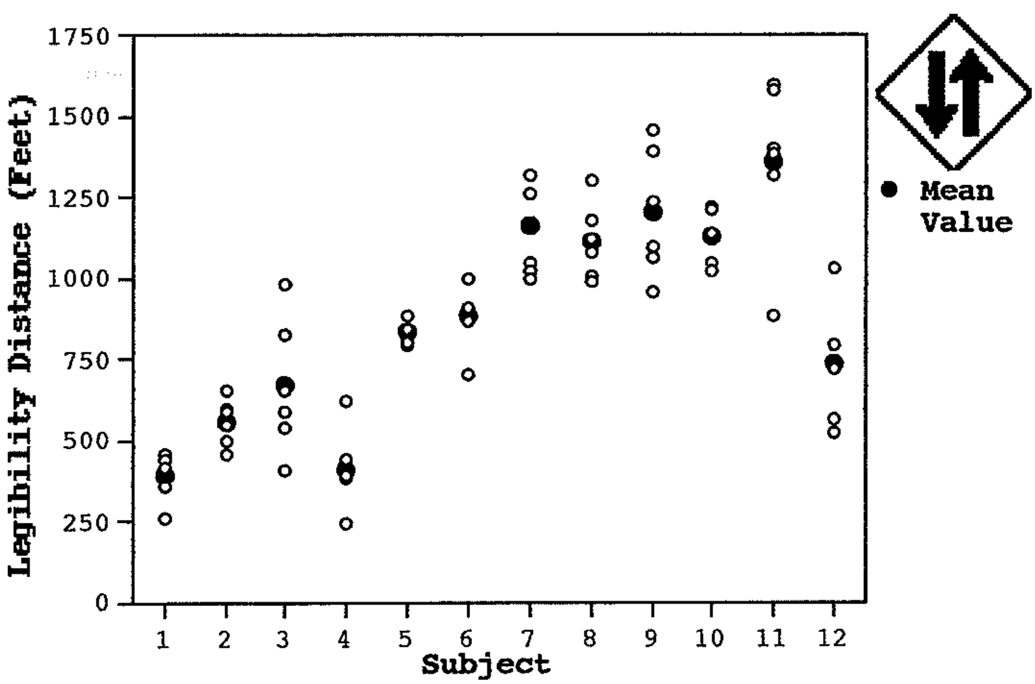

Figure 48. *Laboratory data plot of six trials, with mean. Older drivers and Two-Way Traffic (W6-3) warning sign*

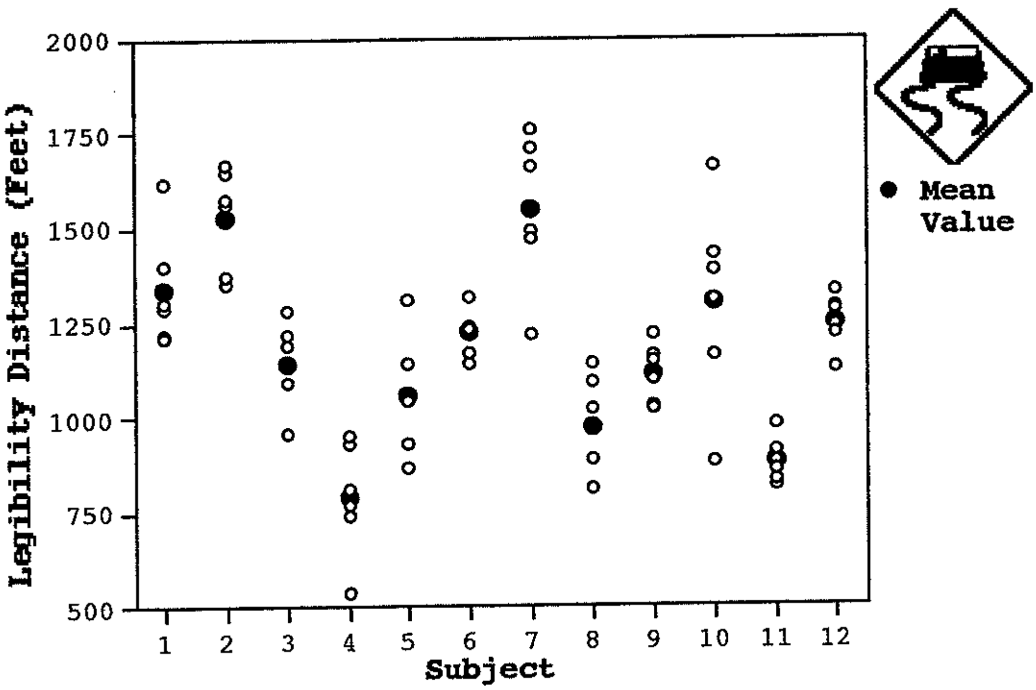

Figure 49. *Laboratory data plot of six trials, with mean. Young drivers and Slippery When Wet (W8-5) warning sign*

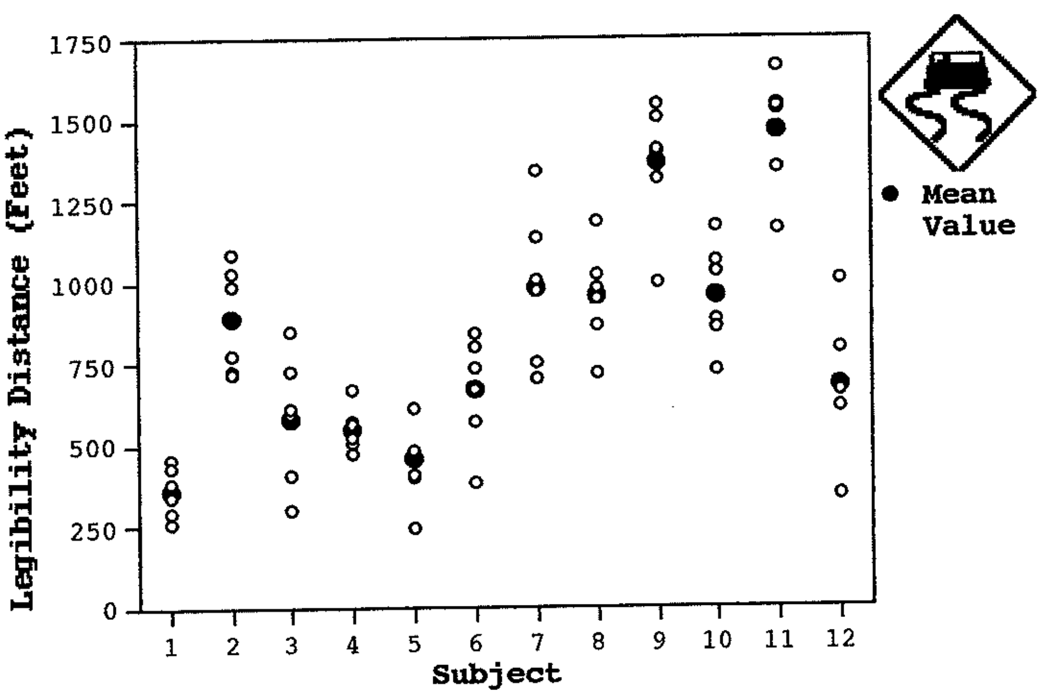

Figure 50. *Laboratory data plot of six trials, with mean. Older drivers and Slippery When Wet (W8-5) warning sign*

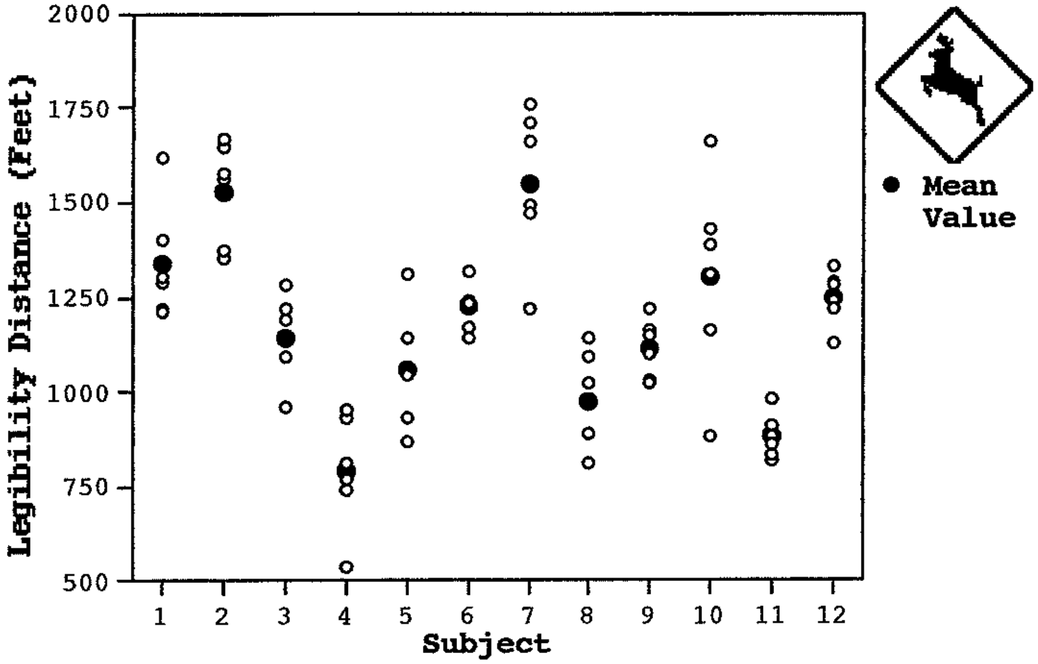

Figure 51. *Laboratory data plot of six trials, with mean. Young drivers and Deer Crossing (W11-3) warning sign*

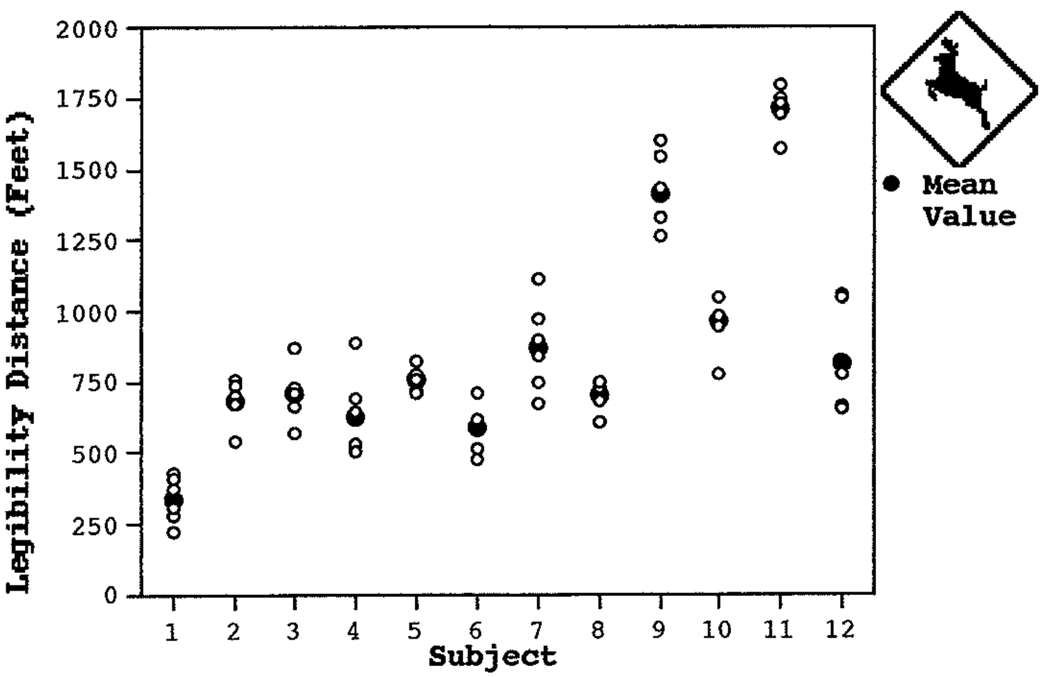

Figure 52. *Laboratory data plot of six trials, with mean. Older drivers and Deer Crossing (W11-3) warning sign*

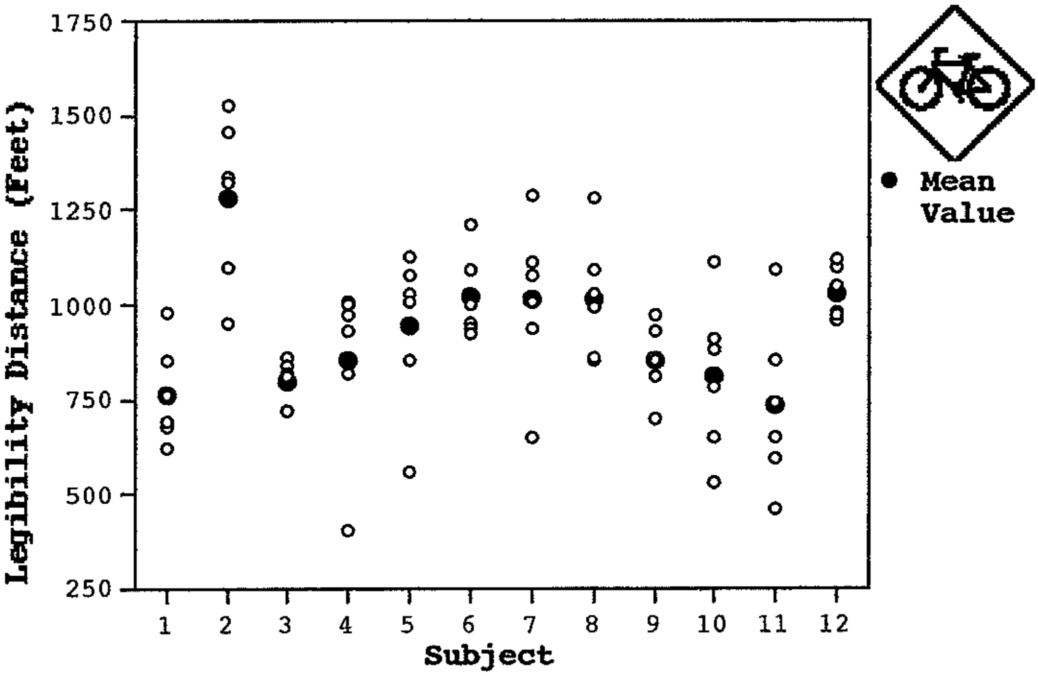

Figure 53. *Laboratory data plot of six trials, with mean. Young drivers and Bicycle Crossing (W11-1) warning sign*

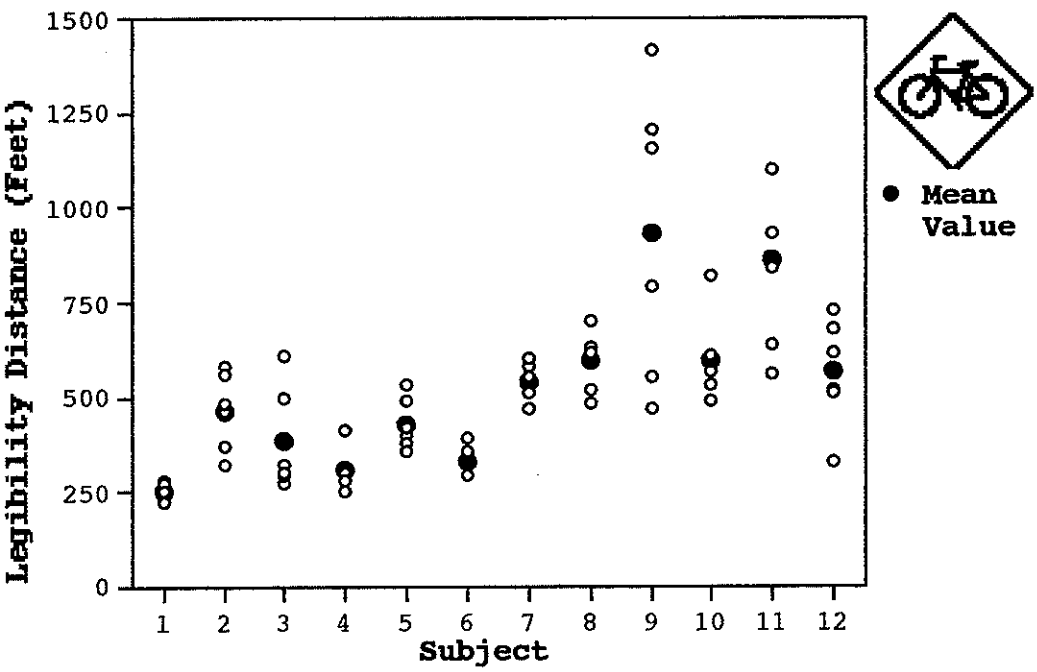

Figure 54. *Laboratory data plot of six trials, with mean. Older drivers and Bicycle Crossing (W11-1) warning sign*

TABLE 30

Ranges of Coefficient of Variation (CoV) by sign, older drivers, laboratory data only

Warning Sign	Minimum CoV	Maximum CoV
Crossroad (W2-1)	3.8	19.9
T-Junction (W2-4)	4.3	33.2
Two-Way Traffic (W6-3)	4.6	31.0
Deer Crossing (W11-3)	4.6	24.4
Bicycle Crossing (W11-1)	9.1	41.4
Slippery When Wet (W8-5)	12.3	34.6

The older subjects are more variable in their responses than the younger drivers. However, it should be noted that even the younger participants are still very variable in their legibility distance responses, as reflected by the large CoV percentages reported in Table 29.

Table 31 is a comparison of the mean CoV by sign for the laboratory data from Bowen (1993) and this laboratory study.

TABLE 31

Comparison of Coefficients of Variation for six warning signs, Bowen's (1993) and present study

Warning Sign	Mean CoV - Bowen (1993)	Mean CoV- This Study
Two-Way Traffic (W6-3)	24.4	14.4
Slippery When Wet (W8-5)	24.5	16.6
T-Junction (W2-4)	30.9	12.9
Crossroad (W2-1)	34.8	11.8
Bicycle Crossing (W11-1)	37.1	19.3
Deer Crossing (W11-3)	40.3	14.4

Bowen's subjects, who were all under the age of 60, were considerably more variable than the subjects who participated in this laboratory study. Bowen used black and white high-contrast symbol images on white cards which started out too small to see and the subject moved the card toward him/her until it was legible. The range of differences between the two studies' CoV finds that Bowen's subjects are between 1.5 and 3.6 times more variable then participants in this laboratory study. It is interesting that the sign with the second largest difference in CoV value is for *Crossroad*, a sign which had some of the lowest CoV values for both field and laboratory sessions. Subjects in Bowen's study were most variable in their legibility response to the sign *Deer Crossing*. An unusual finding was that *Two-Way Traffic*, was the symbol sign with the lowest CoV value for Bowen's study, not *Crossroad*.

COMPARISON OF VARIABILITY BETWEEN FIELD AND LABORATORY DATA

Using the CoV from each of the two studies as the dependent variable, a Pearson product-moment correlation of y's was performed, where y_i was the CoV for field data and y_j was the CoV for laboratory data. The resulting correlation coefficients, r_{ij}, for each warning sign summarizes the strength of the linear relationships between the two response variables and are shown in Table 32.

TABLE 32

Correlation Coefficients for CoV by individual sign

Warning Sign and MUTCD Designation	Correlation Coefficient, r_{ij}
Deer Crossing (W11-3)	0.1583
Bicycle Crossing (W11-1)	0.3457
Two-Way Traffic (W6-3)	0.3915
Slippery When Wet (W8-5)	0.4102
Crossroad (W2-1)	0.5111
T-Junction (W2-4)	0.5446

None of the signs exhibit any strength of relationship between the variability between conditions, field or laboratory. The smallest correlation coefficient (0.1583) is for the *Deer Crossing* sign, and the largest (0.5446) is for *T-Junction*. Large variability by sign for the field study does not imply the same large variability for the laboratory study data. Large within subject variability is present in both conditions and across all signs, but in differing amounts and that variability in one condition cannot predict the variability in the other.

CORRELATION OF MEAN LEGIBILITY DISTANCES BETWEEN CONDITIONS, FIELD AND LABORATORY

Since the original impetus behind the design of this research was to develop a laboratory tool which could be used to predict field legibility distances for new traffic signs, some measure of the correlation between the mean legibility distances for the six signs was calculated. As above, a Pearson product-moment correlation was calculated for each warning symbol sign. The y_i was this analysis was the field mean legibility distance. The y_j was the average legibility distance response for the laboratory study. The computed correlation coefficient, r_{ij}, for each warning sign is presented in Table 33.

TABLE 33

Correlation Coefficients for mean legibility distance, field and laboratory data

Warning Sign	**Correlation Coefficient, r_{ij}**
Two-Way Traffic (W6-3)	0.8195
Bicycle Crossing (W11-1)	0.6484
Crossroad (W2-1)	0.8340
Deer Crossing (W11-3)	0.7822
Slippery When Wet (W8-5)	0.7711
T-Junction (W2-4)	0.8359

The larger r_{ij} values for predicting laboratory legibility distance from field response variables are encouraging. The value of r_{ij} expresses the degree of relationship between the two response variables: field legibility distance and laboratory legibility distance for the six warning symbol signs. The larger values of r_{ij} for the warning symbol sings *Two-Way Traffic*, *Crossroad*, and *T-Junction*, indicate that the laboratory data is a good predictor of field legibility distance. However, that statement is not true for *Bicycle Crossing*, as it has a lower value of r_{ij}, too low perhaps for a traffic engineer to accept the prediction.

Given the high cost and other uncontrolled extraneous variable problems associated with field studies, the laboratory study alternative would be preferred, especially in light of the relatively high positive correlations between the two legibility distances.

ANALYSES INVOLVING SUBSETS OF RESPONSE VARIABLE, LEGIBILITY DISTANCE

Due to the tremendous amount of within subject variability found in the legibility distances in all pilot studies, as well as both the field and laboratory studies, a sense of the impact of this variation was needed. In order to appreciate the within subject variability, subsets of the six replications were taken. The first subset constructed was for the first response given only. This single response was analogous to Paniati's (1988) study of legibility distances of warning signs. The single response is also like the design used by Kline et al. (1990). Tables 34 and 35 summarize, by age group, the mean and standard deviation for legibility distance for all six warning signs, where the data set is composed of the first response given only.

The second subset derived used the first two responses of the six. This set of data was comparable to the two replications collected by Zwahlen et al. (1991), although their subjects were strictly young drivers. Tables 36 and 37 outline the mean and standard deviation, as a function of age, for the data set consisting of the mean of the first two responses given.

The third set of data used the minimum of all six responses. This is a "worst-case" scenario. In the data collected for both studies, the minimum legibility distance was most often not the first response given by a participant. The young drivers' mean and standard deviation for each warning sign is found in Table 38 (and is plotted in Figures 55 and 57), where the response variable analyzed is the minimum of the six responses. Table 39 contains the older subjects' data (and is also plotted in Figures 56 and 58).

As a means of recapping the differences obtained when examining the subsets of data constructed in the same manner as previous experiments, Figure 55 shows the contrast in results for all six symbol signs for the young drivers from the field data. Figure 56 reduces the same data for the older drivers. Figure 57 summarizes the young drivers' results based on the same subsets of response variables from the laboratory data. Figure 58 recapitulates the contrast in results for older drivers. Tabular data used to compile the four Figures, 55 through 58, are contained in Appendix C.

TABLE 34

First response only, young drivers' field and laboratory mean and standard deviation

Warning Sign	Mean Legibility Distance Field*	Standard Deviation Field*	Mean Legibility Distance Laboratory*	Standard Deviation Laboratory*
Two-Way Traffic (W6-3)	341.0 m (1118.7)	93.6 m (307.0)	381.5 m (1251.7)	100.1 m (328.4)
Bicycle Crossing (W11-1)	187.9 m (616.5)	42.4 m (139.1)	266.5 m (874.2)	102.9 m (337.6)
Crossroad (W2-1)	354.1 m (1161.8)	83.1 m (272.6)	429.2 m (1408.3)	81.9 m (268.9)
Deer Crossing (W11-3)	318.2 m (1044.0)	84.1 m (275.9)	356.9 m (1170.8)	99.4 m (326.0)
Slippery When Wet (W8-5)	271.8 m (891.7)	80.1 m (262.9)	342.9 (1125.0)	81.7 m (267.9)
T-Junction (W2-4)	350.9 m (1151.4)	78.4 m (257.1)	434.9 m (1426.7)	94.1 m (308.8)

*** numbers in parentheses denote feet**

TABLE 35

First response only, older drivers' field and laboratory mean and standard deviation

Warning Sign	Mean Legibility Distance Field*	Standard Deviation Field*	Mean Legibility Distance Laboratory*	Standard Deviation Laboratory*
Two-Way Traffic (W6-3)	245.3 m (804.8)	82.4 m (270.2)	242.1 m (794.2)	105.1 m (344.9)
Bicycle Crossing (W11-1)	123.5 m (405.3)	56.2 m (184.5)	133.4 m (437.5)	48.7 m (159.8)
Crossroad (W2-1)	283.7 m (930.7)	102.5 m (336.4)	305.0 m (1000.8)	102.5 m (336.3)
Deer Crossing (W11-3)	166.2 m (545.3)	98.2 m (322.3)	262.4 m (860.8)	106.8 m (350.4)
Slippery When Wet (W8-5)	139.8 m (458.8)	71.4 m (234.1)	213.4 m (700.0)	85.6 m (280.9)
T-Junction (W2-4)	271.5 m (890.6)	131.1 m (430.1)	302.5 m (992.5)	90.0 m (295.4)

* **numbers in parentheses denote feet**

TABLE 36

Mean of first two responses only, young drivers' field and laboratory mean and standard deviation

Warning Sign	Mean Legibility Distance Field*	Standard Deviation Field*	Mean Legibility Distance Laboratory*	Standard Deviation Laboratory*
Two-Way Traffic (W6-3)	343.1 m (1125.5)	94.9 m (311.5)	371.0 m (1217.1)	88.1 m (288.9)
Bicycle Crossing (W11-1)	201.2 m (660.1)	62.8 m (205.9)	263.1 m (863.3)	84.2 m (276.4)
Crossroad (W2-1)	362.6 m (1189.6)	75.4 m (247.5)	415.4 m (1362.9)	82.6 m (271.1)
Deer Crossing (W11-3)	307.2 m (1008.0)	89.1 m (292.2)	337.8 m (1108.3)	83.9 m (275.4)
Slippery When Wet (W8-5)	278.6 m (913.9)	75.8 m (248.8)	352.2 m (1155.4)	74.6 m (244.7)
T-Junction (W2-4)	359.2 m (1178.6)	88.6 m (290.7)	436.6 m (1432.5)	90.4 m (296.5)

* **numbers in parentheses denote feet**

TABLE 37

Mean of first two responses only, older drivers' field and laboratory mean and standard deviation

Warning Sign	Mean Legibility Distance Field*	Standard Deviation Field*	Mean Legibility Distance Laboratory*	Standard Deviation Laboratory*
Two-Way Traffic (W6-3)	244.8 m (803.3)	75.4 m (247.4)	250.9 m (823.3)	101.9 m (334.3)
Bicycle Crossing (W11-1)	128.2 m (420.5)	52.9 m (173.7)	136.4 m (447.5)	52.9 m (173.4)
Crossroad (W2-1)	287.3 m (942.5)	98.4 m (322.9)	311.7 m (1022.5)	109.8 m (360.1)
Deer Crossing (W11-3)	169.9 m (557.3)	88.5 m (290.5)	254.1 m (833.8)	106.1 m (348.1)
Slippery When Wet (W8-5)	157.9 m (518.0)	90.8 m (298.0)	226.3 m (742.5)	100.1 m (328.3)
T-Junction (W2-4)	275.1 m (902.4)	123.8 m (406.2)	316.0 m (1036.7)	96.0 m (315.1)

*** numbers in parentheses denote feet**

TABLE 38

Minimum of six responses, young drivers' field and laboratory mean and standard deviation

Warning Sign	Mean Legibility Distance Field*	Standard Deviation Field*	Mean Legibility Distance Laboratory*	Standard Deviation Laboratory*
Two-Way Traffic (W6-3)	304.9 m (1000.4)	77.9 m (255.6)	326.9 m (1072.5)	74.1 m (243.2)
Bicycle Crossing (W11-1)	178.1 m (584.3)	41.0 m (134.5)	211.3 m (693.3)	58.6 m (192.1)
Crossroad (W2-1)	341.8 m (1121.3)	85.5 m (280.4)	390.9 m (1282.5)	73.5 m (241.2)
Deer Crossing (W11-3)	268.2 m (879.8)	61.9 m (203.0)	303.5 m (995.8)	69.1 m (226.6)
Slippery When Wet (W8-5)	248.0 m (813.8)	57.8 m (189.7)	288.9 m (947.8)	96.7 m (317.1)
T-Junction (W2-4)	325.8 m (1068.9)	82.0 m (268.9)	392.2 m (1286.9)	111.0 m (364.3)

*** numbers in parentheses denote feet**

TABLE 39

Minimum of six responses, older drivers' field and laboratory mean and standard deviation

Warning Sign	Mean Legibility Distance Field*	Standard Deviation Field*	Mean Legibility Distance Laboratory*	Standard Deviation Laboratory*
Two-Way Traffic (W6-3)	214.3 m (703.2)	75.0 m (246.2)	209.0 m (685.8)	90.1 m (295.7)
Bicycle Crossing (W11-1)	117.3 m (384.9)	55.4 m (181.9)	116.6 m (382.5)	38.4 m (126.1)
Crossroad (W2-1)	252.5 m (828.3)	118.4 m (388.5)	280.9 m (921.7)	106.9 m (350.7)
Deer Crossing (W11-3)	139.7 m (458.4)	81.0 m (265.8)	215.6 m (707.5)	110.8 m (363.4)
Slippery When Wet (W8-5)	125.0 m (410.0)	70.3 m (230.5)	149.6 m (490.8)	64.1 m (210.4)
T-Junction (W2-4)	246.8 m (809.8)	123.1 m (403.8)	279.3 m (916.2)	126.0 m (413.3)

* **numbers in parentheses denote feet**

The data depicted in Figures 55 through 58 dramatize the differences in conclusions which would be drawn for legibility distances of the six warning symbol signs. The data in the four columns for each warning sign represent determinations one would make for recommended legibility distances for the six signs in the following manner:

Column 1 - First Response Only - Paniati (1988) and any study which included no replications

Column 2 - Mean of First Two Responses - Zwahlen et al. (1991), since he collected two data points on each of his young subjects.

Column 3 - Worst-case scenario. This minimum of six responses is the shortest legibility distance and is an extremely conservative recommendation. These signs would be legible to a representative sample of older drivers (over the age of 65) or anyone with vision of less than 6.1/12.2 (20/40).

Column 4 - The recommendations which would have followed from these studies, had the within subject variability not been an overshadowing concern and topic for investigation statistically. These distances are in all cases the best or the farthest.

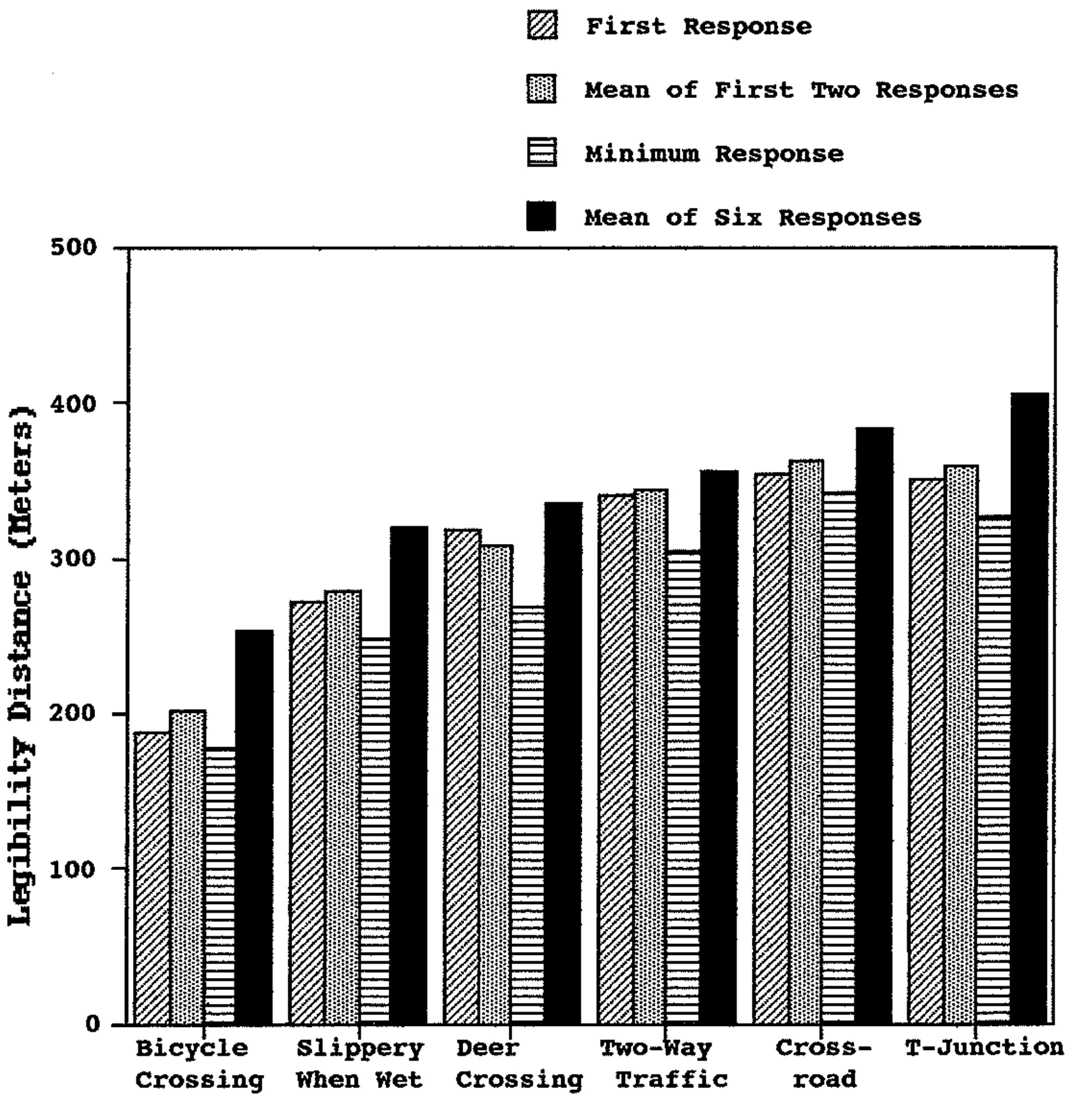

Figure 55. *Summary data plot. Young drivers, field data only, by data subset*

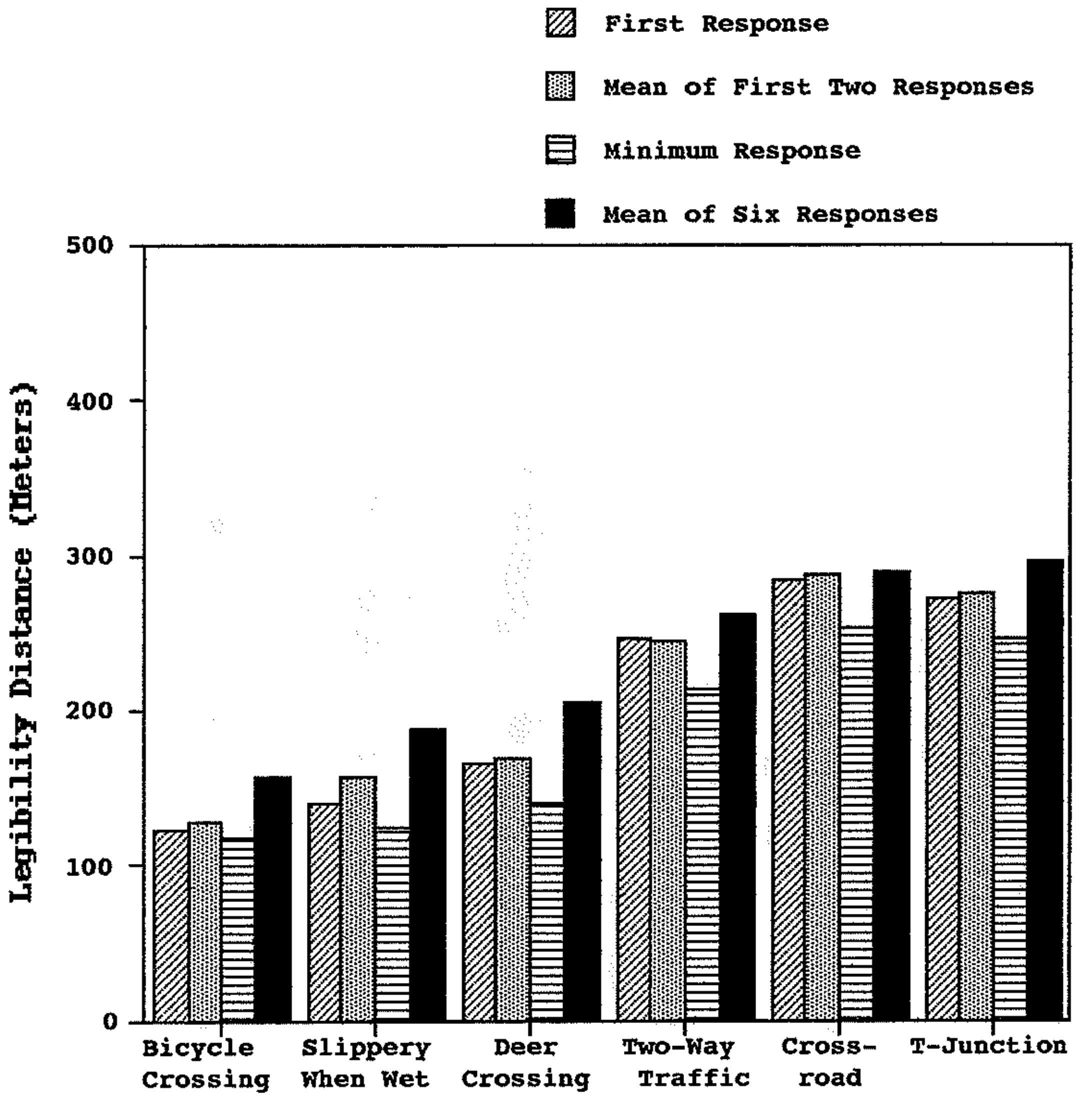

Figure 56. *Summary data plot. Older drivers, field data only, by data subset*

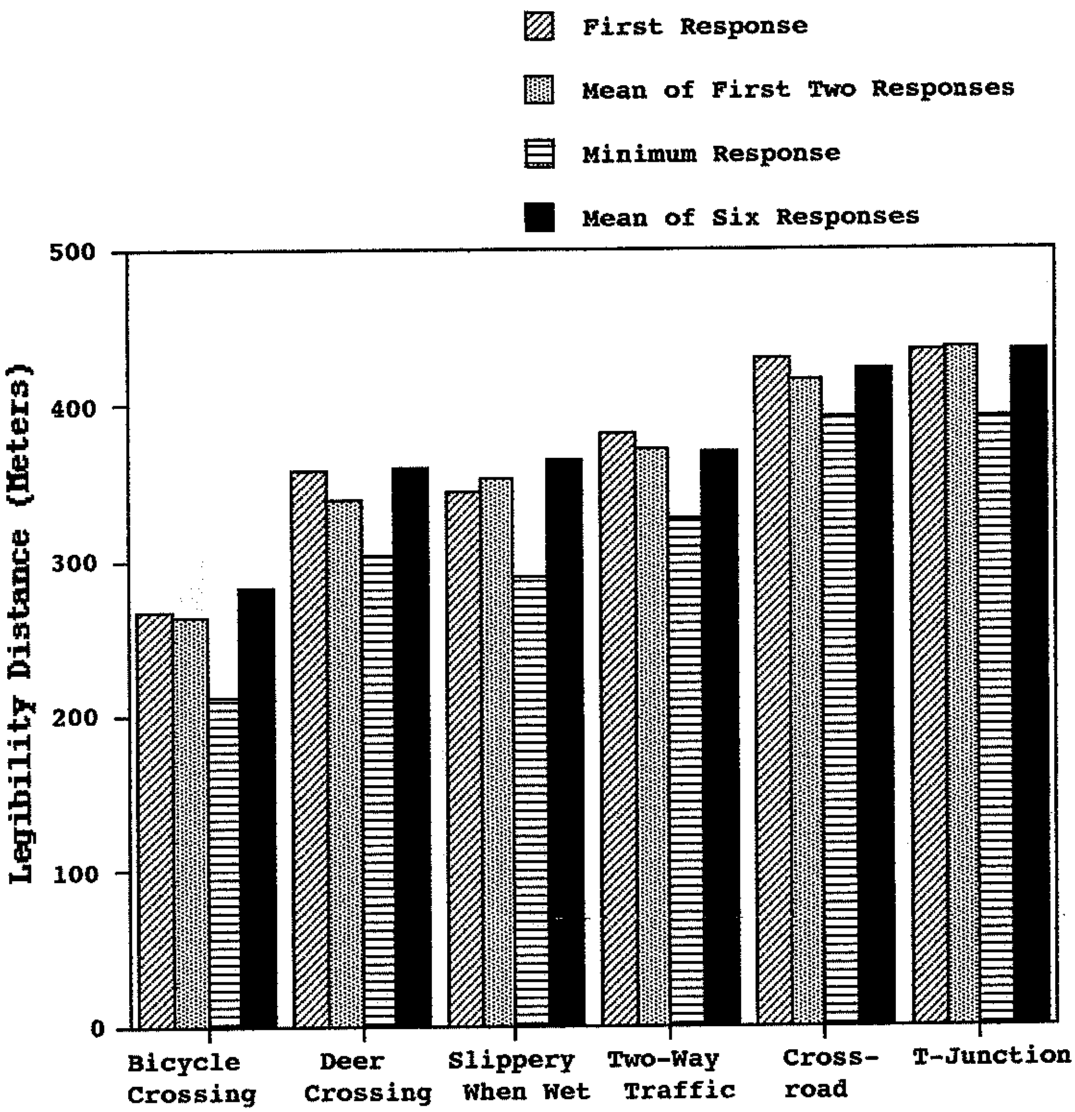

Figure 57. *Summary data plot. Young drivers, laboratory data only, by data subset*

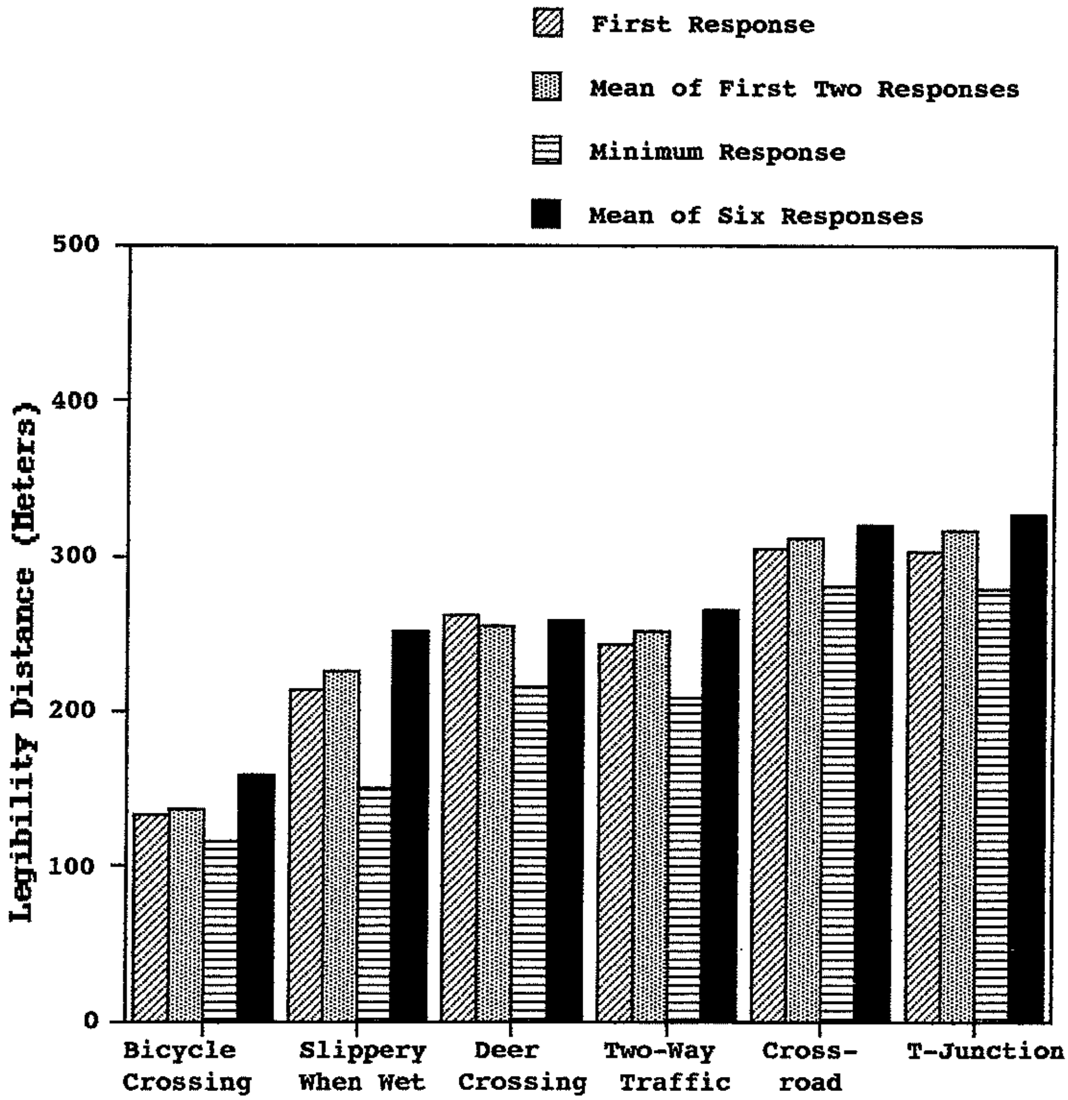

Figure 58. *Summary data plot. Older drivers, laboratory data only, by data subset*

In all cases, of course, the minimum of all responses represents the shortest legibility distance for all signs by age group and condition. This is, as stated above, equivalent to what will be termed the "worst-case" scenario.

For field data, young drivers, the mean of all six responses overestimated the legibility distance for all signs, when compared to the other three subsets of that data. The same observation can be made for older drivers in the field study, with the possible exception of *Crossroad*, where the mean of all six responses is fairly close to the first two columns - first response and mean of the first two responses.

For laboratory data, however, for younger drivers, the mean of the six responses is near the subsets of first response and mean of first two responses. *Bicycle Crossing* is a notable exception. But when older drivers' data is examined for the laboratory, only legibility distance for *Deer Crossing* can be approximated to the mean of the six responses by a subset of the six responses. The legibility distance for the other signs are all overestimated by using the mean of the six responses.

CONCLUSIONS AND RECOMMENDATIONS

The most valuable contribution of this research is the identification of inherent within-subject variability in sign research. The literature is loaded with answers for "predicted legibility distances." But as Zwahlen et al. (1991) point out, those distances vary considerably. They feel the variance is due to setting chosen for the experiment - field or laboratory. The authors postulate several reasons why greater legibility distances are found in the field than the laboratory. However, the present study's legibility distances from the laboratory were greater than the field for all signs and all age groups.

This study was the first of its kind, using a repeated measures design whereby the same young and older subjects were used in both a field and laboratory setting. The same warning symbol signs were investigated in both cases. Replications were used. Six measures of legibility distance on each of six symbol signs were collected from each subject. The older driver population had no one under the age of 65, and spanned the age range through 82.

Another enormous contribution this research has made to the transportation literature is the successful development and proven use of a new laboratory simulation technique. This technique was proven by the high correlation coefficients between laboratory and field legibility distances for all signs. Zwahlen et al.'s (1991) criticism of laboratory studies was valid. Laboratory studies conducted before this new technique was developed were lacking in real-world fidelity. Zwahlen et al. were comparing apples with oranges when examining previous laboratory methods versus the real-world field studies. The prototype simulation developed for this research is a promising first step toward maximizing the parameters which may have caused the dissimilarity in results between prior field and laboratory studies.

The next section will compare the results from both phases of this research with those from past research studies which examined the same signs. Next, an examination of the coefficient of variation and the subsets of data composed from the six responses will be summarized.

Finally, the question will be asked whether or not researchers are worrying about the wrong issue - not where to conduct a study, in the field or laboratory, but what type of subjects should be used and the invaluable collection of multiple observations from the same subject.

COMPARISON OF PRESENT RESULTS WITH PAST STUDIES

Paniati (1988) used both old and young subjects to collect legibility distances for exactly the same six warning symbol signs. It was a laboratory experiment, involving the use of high-quality color slides. The size of the sign started out very small and grew to a size where the subject "identified" it and described its features. As a means of comparison with Paniati's experiment, only the subset "First response only" was used from the laboratory study of this research. Paniati used no replications, just a single response was collected from each subject. Table 40 summarizes the data for both old and young drivers.

TABLE 40

Comparison of mean legibility distance for all signs, Paniati and present research findings

Warning Sign	Paniati Mean Legibility Distance Young	First Response Present Study Young	Paniati Mean Legibility Distance Older	First Response Present Study Older
Deer Crossing (W11-3)	54 m	356.9 m	44 m	262.4 m
Slippery When Wet (W8-5)	76 m	342.9 m	55 m	213.4 m
Bicycle Crossing (W11-1)	148 m	266.5 m	78 m	133.4 m
Two-Way Traffic (W6-3)	150 m	381.5 m	134 m	242.1 m
T-Junction (W2-4)	228 m	434.9 m	128 m	302.5 m
Crossroad (W2-1)	248 m	429.2 m	181 m	305.0 m

The legibility distances obtained in this research, in the laboratory portion, greatly exceed the same distances found by Paniati. Table 41 compares, by age group, the percentage increase this research found over Paniati's legibility distances, by sign.

TABLE 41

Percentage increase comparison of legibility distances, Paniati versus this research

Warning Sign	Percentage Increase in Legibility Distance Over Paniati's Data Young Drivers	Percentage Increase in Legibility Distance Over Paniati's Data Older Drivers
Crossroad (W2-1)	173%	169%
Bicycle Crossing (W11-1)	180%	171%
T-Junction (W2-4)	191%	236%
Two-Way Traffic (W6-3)	254%	181%
Slippery When Wet (W8-5)	451%	388%
Deer Crossing (W11-3)	661%	596%

Even for young drivers, the current study's results would predict a longer legibility distance for all symbols signs, ranging from 1.73 to 6.6 times farther than Paniati would recommend. Similar comparisons exist for older drivers. Here the range of increased legibility distances is 1.69 to 5.96 times farther. For both young and older drivers, *Deer Crossing* stands out as being the sign with the most disparity in results. It is hypothesized that the differences between Paniati's and the present data are directly related to the response variable criterion. For Paniati, the subjects saw each sign only once. Their task was to identify and to describe the features of the sign. *Deer Crossing* is certainly a difficult sign to describe, especially when it might be confused with other signs like *Cattle Crossing* or even *Left Curve*, when seen from a distance.

Bowen (1993) used a laboratory setting, using the same six signs as independent variables. The exact same response variable was collected, legibility distance. The same menu matching technique was also employed between Bowen and this study. Bowen collected five trials on each sign from each subject. The subset closest to Bowen's study is the laboratory study, young driver data only, using all six responses. Table 42 shows that comparison in data results with Bowen's.

TABLE 42

Comparison of mean legibility distance for all signs, Bowen and present research findings

Warning Sign	Mean of Six Responses Bowen Young Drivers	Mean of Six Responses Present Study Young Drivers
Two-Way Traffic - W6-3	418.2 m	369.8 m
Bicycle Crossing - W11-1	328.4 m	282.2 m
Crossroad - W2-1	633.9 m	422.9 m
Deer Crossing - W11-3	499.6 m	359.4 m
Slippery When Wet - W8-5	403.6 m	364.8 m
T-Junction - W2-4	643.1 m	434.7 m

For all signs, Bowen predicts longer mean legibility distances than those found in this study. Table 43 compares the percentage decrease this research found when compared to Bowen's legibility distances, by sign. Only the young driver groups were comparable, therefore, no data exists for comparison with older driver data.

TABLE 43

Percentage comparison of legibility distances between Bowen's and present data

Warning Sign	Percentage of Legibility Distance When Current Data Compared to Bowen - Young Drivers
Crossroad (W2-1)	67%
T-Junction (W2-4)	68%
Deer Crossing (W11-3)	72%
Bicycle Crossing (W11-1)	86%
Slippery When Wet (W8-5)	87%
Two-Way Traffic (W6-3)	88%

In sharp contrast to Paniati's results, Bowen's results overestimate the legibility distances reported in this study. One possible explanation for the differences in results is that Bowen's subjects saw a high-contrast black image on a white background. The current study used the colored signs embedded in a green, treed background. Bowen's conditions represent a baseline, best-case for legibility distance. This research had the color contrast as an added variable, as well as the background, and that resultant contrast.

Zwahlen et al. (1991) conducted a field study using only young drivers. Five of the six signs investigated in the present research were also studied by Zwahlen et al. In a similar fashion, the results from this research's young driver, field condition will be compared in Table 44 with Zwahlen's. The subset of present data chosen for comparison is the mean of the first two responses. This subset is appropriate, since Zwahlen gathered two observations per sign per subject.

TABLE 44

Comparison of mean legibility distance for five signs, Zwahlen et al. (1991) and present research findings

Warning Sign	Mean of First Two Responses Zwahlen et al. Young Drivers	Mean of First Two Responses Present Study Young Drivers
Bicycle Crossing (W11-1)	213.4 m	201.2 m
Slippery When Wet (W8-5)	221.0 m	278.6 m
T-Junction (W2-4)	358.1 m	359.2 m
Two-Way Traffic (W6-3)	373.4 m	343.1 m
Crossroad (W2-1)	428.2 m	363.6 m

Table 45 compares the percentage difference between these research results when compared with Zwahlen et al.'s legibility distances by sign.

TABLE 45

Percentage comparison of legibility distances between Zwahlen et al. (1991) and present data

Warning Sign and MUTCD Designation	Percentage of Legibility Distance When Current Data Compared to Zwahlen et al. (1991) - Young Drivers
Crossroad (W2-1)	85%
Two-Way Traffic (W6-3)	92%
Bicycle Crossing (W11-1)	95%
T-Junction (W2-4)	100%
Slippery When Wet (W8-5)	126%

Of the three studies against which the present research results were compared, The Zwahlen et al.'s field study compares auspiciously with the field results from this study. In three of the five cases the legibility distances predicted from this study underestimate those from Zwahlen et al.'s experiment. *Slippery When Wet* performs better in this study and has a longer legibility distance by 26%. The legibility distances for *T-Junction* differ by just over one meter between the two studies. Zwahlen et al. collected what was termed the distance at which the subject "recognized" the sign. It is unclear from Zwahlen's write-up the precise instructions given to the participants for their response.

COMPARISON WITH OTHER FIELD AND LABORATORY STUDIES

No further direct comparisons can be made with other research studies. This is due to many factors. The methods used by other authors were different, response variable was not defined in the same manner, no replications were used, all conditions are not reported, different warning symbol signs were investigated and/or different instructions and training criterion were applied.

For laboratory studies, as an addendum to this list of possible factors are the issues raised by Zwahlen et al. To review those, the laboratory methodologies are felt to suffer from lack of real-world fidelity or conditions being manipulated in the laboratory which are not, in actual fact, a part of the experimental design. This second criticism encompasses such things as: use of monochrome monitors, projectors or visual displays to generate target signs and scenes. There are by-products of visual displays to project target sign images which are often troublesome, not reported, controlled for or perhaps the authors are not even cognizant of their existence. These by-products include luminances and contrasts, insufficient display resolution, or image vibrations.

SUMMARY OF STUDY RESULTS

For this study, no large surprises emerged for older drivers. Young drivers could see all signs an average of 1.5 times farther than the older participants. Individual signs were different in how far away they were legible and could accurately be matched against the sign menu. There was no age by sign interaction, which meant that the order of legibility distances for signs was the same regardless of age group.

Looking at the field study results, in order of increasing legibility distance, the signs were:

1. *Bicycle Crossing*
2. *Slippery When Wet*
3. *Deer Crossing*
4. *Two-Way Traffic*
5. *Crossroad*
6. *T-Junction*

Both age groups had large percentages of within subject variability. Again, for the field study, the range ran from 31-48%, with the lowest coefficient of variation belonging to *Crossroad*, while the largest variation was found for *Bicycle Crossing*.

The laboratory study results found that in all cases, for all signs and age groups, the legibility distances were longer than for the field study, although not by a large amount. For young drivers, the laboratory results had legibility distances 1.09 times longer than the same drivers in the field condition. For older drivers, the same comparison was 1.12 times farther. In addition, for the laboratory results overall, young subjects saw the signs 1.41 times farther than older drivers.

On a sign-by-sign basis, the young could see the signs farther than the older participants by the following amounts:

Crossroad	1.32
T-Junction	1.33
Two-Way Traffic	1.39
Deer Crossing	1.40
Slippery When Wet	1.45
Bicycle Crossing	1.78

When ordering the signs in increasing legibility distance, the list is exactly the same for the laboratory as it was for the field study.

The coefficients of variation are very large for the laboratory study also, although both groups of subjects are less variable in the laboratory legibility distance responses than the field study, where the range is 28-43%. *Bicycle Crossing* is still the sign which had the most within subject variability associated with responses, and *T-Junction* was the least variable.

Comparing variability between age groups by individual warning sign for the laboratory portion, the numbers are rather surprising. The mean coefficient of variation by sign for older driver compared to the young participants have the following values:

Crossroad	164%
T-Junction	179%
Two-Way Traffic	180%
Deer Crossing	200%
Bicycle Crossing	201%
Slippery When Wet	207%

Older drivers are roughly twice as variable in their responses to three signs, when compared with young drivers in the laboratory environment of this study.

CONCLUSIONS

A very successful new method was developed for a laboratory presentation of target sign stimuli for the purposes of predicting legibility distances for warning symbol signs. This method was based on taking every precaution to use the best equipment and specify every variable not under immediate control. The simulation technique proved to be a valuable contribution to methodologies by which to improve laboratory simulations.

A first of its kind study was conducted using both field and laboratory conditions in a repeated measures design. An equal number of old and young subjects were used. The older subjects were in the age range of 65-82 years of age. A set of baseline legibility distances were collected on the six signs of interest on the same subjects used in the laboratory study.

The purpose behind this baseline establishment was the development of a design support tool in the form of a standardized test and predictive model to be used by traffic engineers. The laboratory data could be used to calculate a regression equation which relates the results from the laboratory to predict field legibility distances. The tool and model could then be validated by using new subjects and new signs and compare model predictions with field observations.

The quandary which occurred was the vast within subject variability obtained in the field and laboratory study dependent variable, legibility distance. The research emphasis turned to the study of this within subject variation and its impact on conclusions drawn.

As seen above, the legibility distances to be recommended for each sign vary depending upon which set of data is examined. The full set of six replications, the first response only, the mean of the first two responses or the minimum of all six responses yield different answers. The contrast in impact these subsets have on legibility distances are a function of age. The older drivers' legibility distances change the most, depending upon which subset is used.

The subset is completely a function of the experimental design and whether or not replications are included. If not, one answer emerges. If two replications are collected, a different answer is given. So, what is the "proper" or "correct" experimental design to use?

Zwahlen et al. make quite a case for performing only field studies as they contain, the authors say, the realism necessary to generate the best or correct answers. They dismiss all laboratory studies as being plagued with all sorts of problems with equipment and specifications and variables which are not controlled. These laboratory *faux pas* are what Zwahlen et al. feel account for the differences in legibility distances for field and laboratory studies. According to Zwahlen et al. the laboratory studies underestimate legibility distances by a factor of 1.7, when compared to field data.

More than just experimental setting is causing the differences in legibility distances. Field studies are not without their drawbacks. Besides being costly and time-consuming, they can put the driver in a dangerous situation. Uncontrollable extraneous variables run rampant in field research.

Zwahlen et al. are asking the research community to make a change to field studies only, or be able to better specify and control laboratory variables. This is the wrong plea to be making to the transportation research community. Instead, the laboratory simulation developed for this research shows that laboratory methodologies do work and can predict field legibility distances with good accuracy.

RECOMMENDATIONS

It is recommended that the area of within subject variability in legibility distance responses be given serious examination and recognition for its existence. Also, the ramifications this variance has on predictions must be reconciled.

Researchers must come to grips with three issues of foremost importance: first, a representative sample of older drivers must be included in every study. Secondly, the use of a larger sample size than which generally appears in the literature is also called for. The literature supports time and time again problems the aging process has on the visual system and reaction time. To continue the practice of using only young subjects to generate recommendations for sign placement is senseless and ignores an important and growing segment of our population, the older driver.

Finally, replications are vital and must be included in every study, whether it is field- or laboratory-based. Without replications, we, as practicing transportation researchers, are fooling ourselves if we think we can collect just one or maybe two observations and make recommendations. The conclusions in the literature which are based on only one observation, or perhaps two, are very likely misleading and in error.

Without replications, the author is overlooking that large and consequential within subject variation contribution. It is highly recommended the research community recognize this phenomenon. Since without a thorough study of this factor, perhaps that which is being reported is more a function of the experimental design than the multitude of independent variables being painstakingly manipulated.

The technology in computers and multi-media has evolved to a point now where it is cheap and available. A standardized laboratory test could easily be developed using video-taped imagery of actual signs embedded in real backgrounds. This approach, although considered, was not feasible due to the equipment availability at the time. The videotaped images are memory intensive, but that can be solved with a laser disc and a random-access laser disc player. The videotape camera issue was examined at length for this study and an affordable alternative identified. The cost to create a laser disc has dropped dramatically in price recently from the $3,000+ price tag of only a few years ago, to an affordable $250-$350. The laser disc player is also in a price range which is cost-effective as an investment by a transportation district or State DoT headquarters. The cost would not exceed $600. The computers and software, as well as the people with the knowledge to integrate all the hardware, peripherals and software exist. The laboratory prototype is not only realistic, but possible with a minimal dollar investment. The benefits would amortize the equipment quickly.

This study, with its use of a repeated measures design and replications uncovered a new, never-before discussed parameter affecting legibility distance - within subject variability. It is substantial, and can possibly be causing the differences reported in the literature of legibility distances of symbol signs. Recommended distances at which signs are placed to be recognized must be determined from a "worst-case" scenario. This premise requires a reexamination of our research methodologies for determining guidelines for placement of warning symbol signs. The source of this within subject variability is not identified. The variance displayed by the human observer, regardless of age, appears to be a stochastic process. Further research must be directed at distinguishing the source of the variability and a means whereby a correction factor is constructed to account for it.

REFERENCES

Allen, M.J. (1970). *Vision and Highway Safety.* Philadelphia: Chilton Book Company.

Allen, R.W., Parseghian, Z., and Van Valkenburgh, P.G. (1980). *Age effects on symbol sign recognition* (Tech. Report FHWA/RD-80/126). Hawthorne, CA: Systems Technology, Inc.

Anderton, P.J., Johnston, A.W., and Cole, B.L. (1974). The effect of letter spacing on the legibility of direction and information signs. *Australian Road Research, 5*(5), 100-102.

Avant, L.L., Brewer, K.A., Thieman, A.A., and Woodman, W.F. (1977). *Recognition errors among highway signs* (Transportation Research Record 1027 - "Driver Information Needs and Visibility of Traffic Control Devices"), pp. 45-49. Washington, DC: National Research Council, Transportation Research Board.

Blackwell, H.R., and Blackwell, O.M. (1980). Population data for 140 normal 20-30 year olds for use in assessing some effects of lighting upon visual performance. *Journal of the Illuminating Engineering Society, 9*(3), 158-174.

Blackwell, O.M., and Blackwell, H.R. (1980). Individual responses to lighting parameters for a population of 235 observers of varying ages. *Journal of the Illuminating Engineering Society, 9*(6), 205-232.

Brewer, K.A., Thieman, A.A., Woodman, W.F., and Avant, L.L. (1980). *Highway sign meaning as an indicator of perceptual response* (Transportation Research Record 1027- "Driver Information Needs and Visibility of Traffic Control Devices"), pp. 35-43. Washington DC: National Research Council, Transportation Research Board.

Burg, A. (1966). Visual acuity as measured by dynamic and static tests. *Journal of Applied Psychology, 50*, 460-466.

Cooper, B.R. (1988). *A comparison of different ways of increasing traffic sign conspicuity.* (Research Report 157). Crowthorne, England: Traffic Operations Division, Traffic Group, Transport and Road Research Laboratory.

Croxton, F.E., and Cowden, D.J. (1964). *Applied General Statistics* (2nd ed.). New Jersey: Prentice-Hall, Inc.

Desrosiers, R.D. (1965). Moving picture technique for highway signing studies: an investigation of its applicability. *Public Roads, 33*(7), 143-147.

Dewar, R.E. (1972). Permission versus prohibitive symbols for regulatory traffic control signs. In *Proceedings of the International Conference on Highway Sign Symbology* (pp. 20-27). Washington, DC: International Road Federation and U.S. Department of Transportation, Federal Highway Administration.

Dewar, R.E. (1976). The slash obscures the symbol on prohibitive traffic signs. *Human Factors, 18*(3), 253-258.

Dewar, R.E. (1979). *Criteria for the design and evaluation of traffic sign symbols* (Transportation Research Record 1160), pp. 1-6. Washington DC: National Research Council, Transportation Research Board.

Dewar, R.E. (1979) Static traffic signs: state-of-the-art review. In *Proceedings of the International Symposium on Traffic Control Systems. Volume I - Resource papers (UCB-ITS-P79-2)* (pp. 138-166). W.S. Homburger and L. Steinman (Eds.). Contract DOT-FH-119452. Washington, DC: United Stated Department of Transportation, Federal Highway Administration, Office of Research.

Dewar, R.E., and Berger, W. (1972). Methodology in traffic sign evaluation. In *Proceedings of the International Conference on Highway Sign Symbology* (pp. 56-60). Washington, DC: International Road Federation and U.S. Department of Transportation, Federal Highway Administration.

Dewar, R.E., and Ells, J.G. (1974). *Comparison of three methods for evaluating traffic signs*. (Transportation Research Record 503) pp. 38-47. Washington, DC: National Research Council, Transportation Research Board.

Dewar, R.E., Ells J.G., and Mundy, G. (1976). Reaction time as an index of traffic sign perception. *Human Factors, 18*(4), 381-392.

Dewar, R.E., Kosnik, W., and Kline, J. (1992). Driver age and comprehension of symbolic highway signs. Task D, Study 2: Report of Symbol Signing Design for Older Drivers. Washington, DC: National Research Council, Transportation Research Board.

Dewar, R.E., and Swanson, H.A. (1972). *Recognition of traffic-control signs*. (Highway Research Record 414) pp. 16-23. Washington, DC: National Research Council, Transportation Research Board.

Ells, J.G., and Dewar, R.E. (1979). Rapid comprehension of verbal and symbolic traffic sign messages. *Human Factors, 21*(2), 161-168.

Ericksen, C.W., Hamlin, R.M., and Breitmeyer, R.G. (1970). Temporal factors in visual perception as related to aging. *Perception and Psychophysics, 7*,(6), 354-356.

Evans, D.A., and Ginsburg, A.P. (1985). Contrast sensitivity predicts age-related differences in highway-sign discriminability. *Human Factors*, 27(6), 637-642.

Federal Highway Administration (1983). *Traffic Control Devices Handbook*. Washington, DC: U.S. Department of Transportation.

Federal Highway Administration (1988). *Manual on Uniform Traffic Control Devices for streets and highways*. Washington, DC: U.S. Department of Transportation.

Forbes, T.W. (1939). A method for analysis of the effectiveness of highway signs. *Journal of Applied Psychology, 23*, 669-684.

Forbes, T.W. (1964). Predicting attention-gaining characteristics of highway traffic signs: measurement technique. *Human Factors, 6*, 371-374.

Forbes, T.W. (1972). Visibility and legibility of highway signs. In T.W. Forbes (ed.), *Human factors in highway traffic safety research* (pp. 95-109). New York: Wiley-Interscience, John Wiley.

Forbes, T.W. (1976). *Luminance and contrast for sign legibility and color recognition*. (Transportation Research Record 611), pp. 17-24. Washington, DC: National Research Council, Transportation Research Board.

Forbes, T.W. (1980). *Acuity, luminance and contrast for highway sign legibility: samples of research methods and results. A review of fifteen selected studies by various investigators.* East Lansing, MI: Michigan State University, Department of Psychology and Highway Traffic Safety Center.

Forbes, T.W., Saari, B.B., Greenwood, W.H., Goldblatt, J.G., and Hill, T.E. (1976). *Luminance and contrast requirements for legibility and visibility of highway signs.* (Transportation Research Record 562), pp. 59-72. Washington, DC: National Research Council, Transportation Research Board.

Forbes, T.W., Snyder, T.E., and Pain, R.F. (1964). A study of traffic sign requirements. Vol. II: An annotated bibliography. East Lansing, MI: College of Engineering, Michigan State University , under contract with Minnesota Mining and Manufacturing Company.

Gibson, J.J. (1979). *The ecological approach to visual perception.* Boston: Houghton-Mifflin.

Giesa, S. (1972). Suggestions for improving the unification of traffic signs and symbols. In *Proceedings of the International Conference on Highway Sign Symbology* (pp. 39-50). Washington, DC: International Road Federation and U.S. Department of Transportation, Federal Highway Administration.

Ginsburg, Arthur P., Evans D. W., Cannon M.W., Owsley C., and Mulvanny, P. (1984). Large-sample norms for contrast sensitivity. *American Journal of Optometry and Physiological Optics, 61*(2), 80-84.

Halpern, D.F. (1984). Age differences in response time to verbal and symbolic traffic signs. *Experimental Aging Research, 10*(4), 201-204.

Hills, B.L. (1972). Measurement of the night-time visibility of signs and delineators on an Australian rural road. *Australian Road Research, 4*(10), 38-57.

Hills, B.L. (1975). Visibility under night driving conditions. Part 1: Laboratory background and theoretical considerations. *Lighting Research and Technology, 7*(3), 179-184.

Hills, B.L. (1975). Visibility under night driving conditions. Part 2: Field measurements using disc obstacles and a pedestrian dummy. *Lighting Research and Technology, 7*(4), 251-258.

Hodge, D.C. (1962). Legibility of a uniform-stroke alphabet: Volume I: Relative legibility of upper and lower case letters. *Journal of Engineering Psychology, 1*, 34-46.

Jacewicz, M.M., and Hartley, A. A. (1987). Age differences in the speed of cognitive operations: resolution of inconsistent findings. *Journal of Gerontology, 42*, 86-88.

Jacobs, R.J., Johnston, A.W., and Cole, B.L. (1975). The visibility of alphabetic and symbolic traffic signs. *Australian Road Research, 5*(7), 68-86.

Kline, D. (1991). *Aging and the visibility of highway signs: a new look through old eyes.* Unpublished Technical Report for the AAA Foundation for Traffic Safety. Calgary, Canada: Vision and Aging Laboratory, Department of Psychology, The University of Calgary.

Kline, T.J.B., Ghali, L.M., and Kline, D. (1990). Visibility distance of highway signs among young, middle-aged, and older observers: icons are better than text. *Human Factors, 32*(5), 609-619.

Kline, D., and Schieber, F. (1982). Visual persistence and temporal resolution. In R. Sekuler, D. Kline, and K. Dismukes (Eds.), *Aging and human visual function* (pp. 231-244). New York: Alan R. Liss.

Knoblaugh, R.L., and Pietrucha, M.T. (1986). *Motorists' comprehension of regulatory, warning and symbol signs: Volume II: Technical report* (FHWA/RD-86/112). Washington, DC: Department of Transportation, Federal Highway Administration.

Kosnick, W.D., Sekuler, R., and Kline, D.W. (1990). Self-reported visual problems of older drivers. *Human Factors, 32*(5), 597-608.

Lum, H.S., Roberts, K.M., DiMarco, R.J., and Allen, R.W. (1983). A highway simulator analysis of background colors for advance warning signs. *Public Roads, 47*(3), 89-96.

Mace, D.J. (1988). Sign legibility and conspicuity. In *Transportation in an aging society.* (Special Report 218, Vol. 2, pp. 270-293). Washington, DC: National Research Council, Transportation Research Board.

Mace, D.J., and Gabel, R. (1992). *Modeling highway visibility minimum required visibility distance.* Presented at the 71st Annual Meeting of the Transportation Research Board, Washington, DC.

Mace, D.J., and Pollack, L. (1983. *Visual complexity and sign brightness in detection and recognition of traffic signs* (Transportation Research Record 904), pp. 33-41. Washington, DC: National Research Council, Transportation Research Board.

Mackie, A.M., and Higgs, M.H. (1972). Motorist's understanding of the meanings of symbolic traffic signs. In *Proceedings of the International Conference on Highway Sign Symbology* (pp. 61-79). Washington, DC: International Road Federation and U.S. Department of Transportation, Federal Highway Administration.

Markowitz, J., Dietrich, C.W., Lees, W.J., and Farman, M. (1968). *An investigation of the design and performance of traffic control devices.* (Contract CPR-11-5955, Report No. 1726). Cambridge, MA: Bolt, Beranek, and Newman, Inc.

National Research Council (1988). *Special Report 218: Transportation in an aging society: improving mobility and safety for older persons, Volume 1.* Washington, DC: Transportation Research Board, U.S. Department of Transportation.

Nelson, M.A., and Halberb, R.L. (1979). Visual contrast sensitivity functions obtained with colored and achromatic gratings. *Human Factors, 21*(2), 225-228.

Olson, P.L., and Bernstein, A. (1979). The nighttime legibility of highway signs as a function of their luminance characteristics. *Human Factors, 21*(2), 145-160.

Paniati, J.F. (1988). Recognition and comprehension of traffic sign symbols. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 568-572). Santa Monica, CA: Human Factors Society.

Planek, T.W., and Overend, R.B. (1973). How aging affects the driver. *Traffic Safety, 73*(2).

Ranney, T.A. and Pulling, N.H. (1990). *Performance differences on driving and laboratory tasks between drivers of different ages* (Transportation Research Record 1281), pp. 3-10. Washington DC: National Research Council, Transportation Research Board.

Schaie, K., Warner, K., and Gribben, K. (1975). Adult development and aging. *Annual Review of Psychology, 26*, 65-96.

Shoptaugh, C.F., and Whitaker, L.A. (1984). Verbal response times to directional traffic signs embedded in photographic street scenes. *Human Factors, 26*(2), 235-244.

Sivak, M., and Olson, P.L. (1981). Effect of driver's age on nighttime legibility of highway signs. *Human Factors, 23*(1), 59-64.

Sivak, M., Olson, P.L., and Pastalan, L.A. (1985). *Optimal and minimal luminance characteristics for retroreflective highway signs* (Transportation Research Record 1027), pp. 53-57. Washington DC: National Research Council, Transportation Research Board.

Walker, R.E., Nicolay, R.C., and Stearns, C.R. (1965). Comparative accuracy of recognizing American and International road signs. *Journal of Applied Psychology, 49*, 322-325.

Whitaker, L.A. (1985). *Driver's unconscious errors in the processing of traffic signs.* (Transportation Research Record 1027), pp. 42-45. Washington DC: National Research Council, Transportation Research Board.

Whitaker, L.A. and Stacey, S. (1981). Response times to left and right directional signs. *Human Factors, 23*(4), 447-452.

Winters, D.J. (1985). Learning and motivational characteristics of older people pertaining to traffic safety. In J.L. Malfetti (Ed.), *Drivers 55+: Needs and problems of older drivers: Survey results and recommendations* (pp. 77-86). Falls Church, VA: AAA Foundation for Traffic Safety.

Yee, D. (1985). A survey of traffic safety needs and problems of drivers age 55 and over. In J.L. Malfetti (Ed.), *Drivers 55+: Needs and problems of older drivers: Survey results and recommendations* (pp. 96-128). Falls Church, VA: AAA Foundation for Traffic Safety.

Zwahlen, H.T., Hu, X., Sunkara, M., and Duffus, L.M. (1991). Recognition of traffic sign symbols in the field during daytime and nighttime. In *Proceedings of the Human Factors Society 35th Annual Meeting* (pp. 1058-1062). Santa Monica, CA: Human Factors Society.

APPENDIX A

CREATION AND EDITING OF SYMBOL IMAGES FOR LABORATORY SIMULATION

Aldus SuperPaint®, Version 3.0 images of the warning symbol signs were bit-mapped into 256 x 256 arrays of pixel elements (Bowen, 1993). Using the detailed drawings of the warning symbols signs from the Standard Symbol Signs (1979) (SSS) Manual as a guide, the symbol signs were built and stored in a PICT file format.

Note that the overall height and width of the symbols and their placement on the signs are provided in the SSS Manual. A grid is supplied to specify the exact detail of the symbol. Figure 59 is a copy of the applicable grid page taken directly from the Manual. Figure 59 shows the grid layout for the *Deer Crossing* sign.

It is obvious from examining the grid in Figure 59 that some features of the sign, especially parts of the deer's legs, tail and antlers, do not encompass one grid square. On the computer monitor available for this research, display resolution is limited to 72 pixels per inch. Pixels can only be square, whereas many parts of complex signs, like *Deer Crossing*, are not. The resulting bit-mapped PICT images of the signs were as close to the shape dictated by the Standard Symbol Sign Manual as possible.

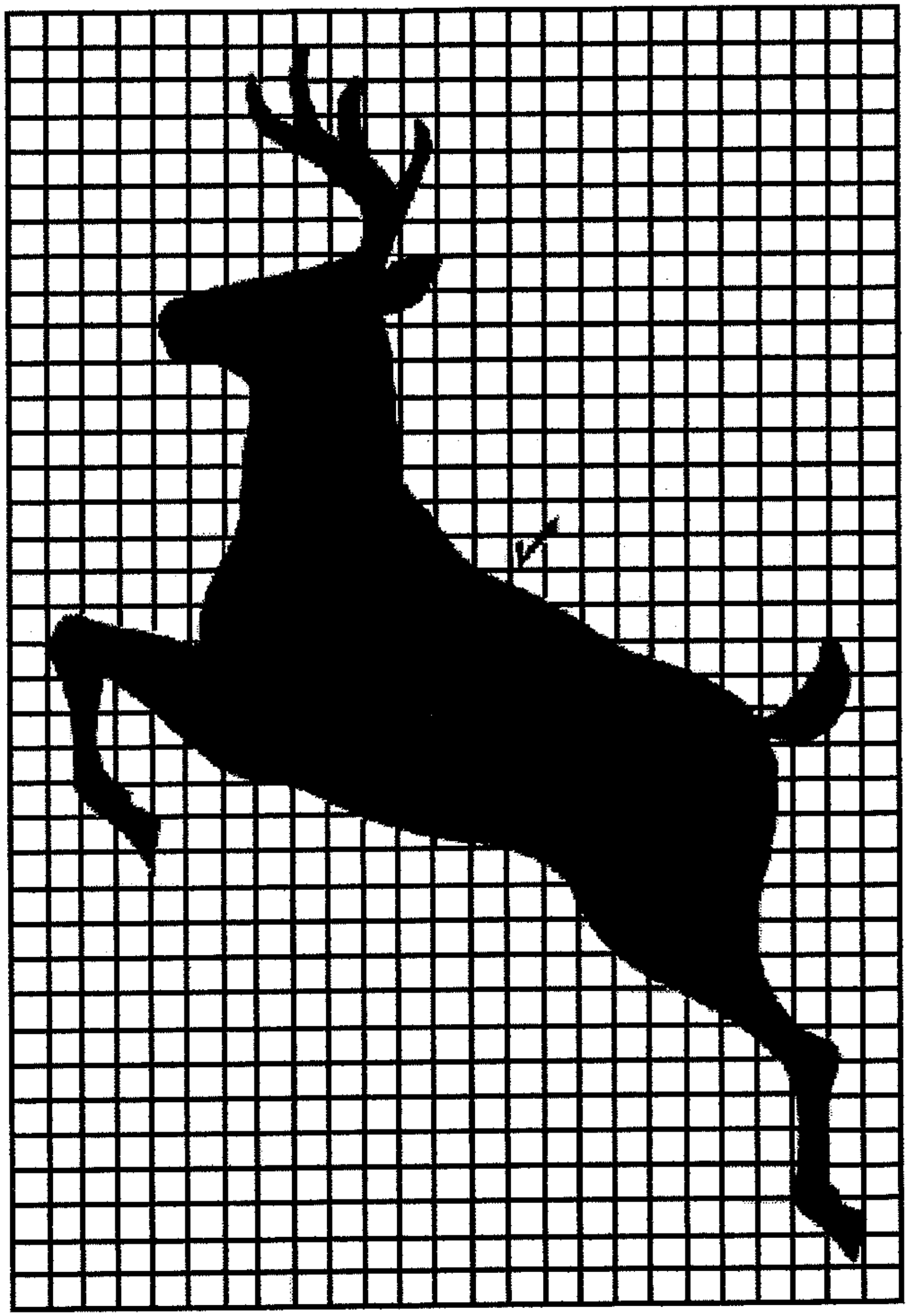

Figure 59. *Page 2-60 from Standard Symbol Signs Manual (1979), grid specifications for Deer Crossing (W11-3)*

The Viewsonic® 7 color monitor used for the laboratory experiment was driven by an 8-bit color video card, which meant it displayed 2^8 or 256 colors. Without the benefit of a spectroradiometer to measure the precise u',v' (Committee Internationale de l'Eclairge, 1976) color coordinates of real-world warning symbol signs, a yellow from the 256-color palette was chosen which was as close to the "yellow" of the actual signs. This match was made subjectively.

It is recognized that in order to have exactly the same color contrast as real-world signs, the color would have to be matched exactly. Matching those color coordinates might have been possible with the aid of two additional pieces of hardware: the spectroradiometer, as already mentioned, plus a 24 bit color graphics video card. This 24-bit card would allow the display of 2^{24} or 16,777,216 colors. Surely, one of those colors would be close to the yellow of the warning signs.

Once a standard 256 x 256 pixel array was created for each sign of interest (the resulting diagonal size of each sign was 8.467 cm or 3.333 inches), a sizing algorithm had to be established, resulting in signs of varying visual angles equivalent to that of full-sized signs viewed from different distances.

As described above, the 256 x 256 pixel array composed each symbol sign and served as the standard. This standard was scaled to the appropriate percentage, based upon the number of pixels required to subtend the visual angle of the sign from approaching distances.

A software program scaled the 256 x 256 standard image. Deneba Canvas™, Version 3.0 read each SuperPaint® PICT file. Scaling the original one-bit depth (black and white) PICT file resolution caused distortion of the sign image. It was found that increasing the bit depth resolution of the PICT files eliminated or minimized the distortion. The image depth was increased from one bit (black and white), in which they were created, to a 32 bit image.

Although the original images were created only in black and white or a one-bit resolution, artificially increasing the depth to 32-bit effectively made the image as if it were composed of 2^{32} shades of black. Along with this bit-mapped image being increased in color depth, the display resolution was also increased.

The monitor resolution used to present the simulation had a 0.28 mm dot pitch, or could display 72 dots per inch (dpi). Although the display resolution of the Viewsonic® monitor is only 72 pixels per inch (equivalent to dots per inch or dpi), by raising the sign image resolution to 300 dpi the result was essentially the equivalent of taking each pixel and splitting it into four pixels for editing purposes. The sign bit-mapped image could be edited at a higher resolution, although it still was only displayed at 72 dpi. This conversion was done to preserve the pixels when the original image was down-scaled. The final step after increasing the image bit depth and display resolution, color was added to the black and white sign image background from the 256 color palette, as discussed above.

By increasing the resolution of the PICT image to 300 dpi, gradients resulted when the image was scaled down using the software. Figure 60 shows for the *Deer Crossing* sign, one scaled at 37.5% and at 8.91% of the original 256 x 256 pixel image, corresponding to distances of 441.9 and 103.6 meters (1450 and 340 feet) from the sign when viewed from 6.1 meters (20 feet) away.

Figure 60. *Scaled PICT images of Deer Crossing sign as sized at 441.9 and 103.6 meters from sign*

When examining Figure 60 without the benefit of a full-color printout, it is not evident where the gradients are in the scaled down image. However, where the background color of yellow pixels met a black pixel of the deer symbol, the result was to create different color pixels. Those pixels were different shades of gray, green and other blendings of yellow and black.

On the color monitor the result of scaling down the black on yellow image was to create an image with gradients of colors. Each pixel was not always a pure black or a pure yellow. A gradient in pixels occurred where the yellow met the black, for instance. Perimeter black pixels were graded into the yellow background. This gradient result was a by-product of the Canvas™ scaling routine.

Moreover, as stated above, since the original 256 x 256 pixel image consisted of square pixels which were either yellow or black, the symbol on the sign was not always exactly the shape as dictated by the Standard Symbol Sign Manual. Down or up scaling the 256 x 256 pixel array did not make those square pixels change shape. So the resultant image was again, a less-than-perfect scaled image of the symbol. It would have been preferable if computer image scaling worked like the percentage reduction or enlargement on a copier machine.

The copier works with a camera lens which can blow up or shrink the image in continuous increments. The computer software operates in the following manner. A 256 x 256 pixel image is created which ends in a 3.3333 inch high by 3.33333 inch wide sign image. This image must be scaled by a factor of 10.44% to produce the size image equivalent of viewing the original 30 inch sign from a distance of 1220 feet. The mathematics would say to simply multiply 256 pixels by 10.44% to find out the resulting size of the scaled image. However, the problem which arose was that the mathematical answer says an image with a height and width of 26.7264 pixels must be produced. The software cannot produce 0.7264 pixel. As a matter of fact, the Canvas™ software does not round up the decimal pixels. Consequently, any answer from 26.00001 to 26.99999 pixels creates an image which is 26 pixels. Integer pixels was another reality which had to be considered in the scaled image development.

Since all signs were scaled from a 256 x 256 pixel array, the same scaling was required for each distance interval. Each sign consisted of two parts - the yellow diamond with the black border and second, the symbol. It was decided to scale the two elements separately and merge them in the animation. This made any "clean-up" tasks a one-time requirement. For instance, as the sign border was scaled down in size, the pixels often needed some editing because of the integer pixel problem. However, the same integer values of pixels were required for each symbol sign diamond. So creating the scaled diamonds and cleaning them up once proved to be a time-saving strategy. A library of individual scaled diamonds, representing each distance viewed, was created.

Next, the symbol element inside the diamond was isolated, with enough of the yellow background pixels to allow for gradients upon scaling. Then each symbol was scaled, using the same procedure and software described above. Another library, one for each sign to be presented in the simulation, was produced.

APPENDIX B

INSTRUCTIONS TO PARTICIPANTS - LABORATORY STUDY

You are about to participate in the second portion of the sign recognition study, this time involving a computer simulation. You will be seated at a table 10 feet away from a large screen, high resolution monitor. The computer will present a static digitized scene from the Annex, where you participated in Phase I, driving the research vehicle. The scene you will see will be similar to the background you encountered at the Research Annex. This background on the computer will NOT move, however, the signs embedded in the background will increase in size and appear to move toward you, as to simulate a driving speed of approximately 20 miles per hour.

In front of you, there is a menu of the 27 possible signs. This is the same menu we reviewed and used in the field study. You must refer to this menu at all times. Using this menu, you will be required to give two responses. If you recall from the field study, the two responses which I am looking for are exactly the same as the previous ones. As in Phase I, please give me a guess as to what you believe the sign to be and then when you are certain. Again, only respond "certain" when you are positive as to the sign's identity and would be willing to make a vehicle maneuver to the sign's content.

Please only respond with the number of the sign. There is no penalty for being wrong. I will not give you any feedback on your guesses. If you believe you are certain about a sign's identification, but are incorrect, I will inform you. At this point, just continue watching the monitor until the sign increases in size to a point where you believe you are certain as to its identity.

When you are correct, I will tell you so. The computer screen will go blank for a few moments while the next sign file is loading in the computer. The next trial is automatically loaded and we will commence another trial session when you signal you are ready.

It is expected that the trials will last approximately one hour. The first two trials, as a minimum, will be used for practice. We will only commence the actual data trials when you feel you are comfortable with the simulation and your task. We will have as many practice trials as you would like before we begin.

Do you have any questions?

I MUST EMPHASIZE YOUR RESPONSE TECHNIQUE. PLEASE ONLY RESPOND BY THE NUMBER OF THE SIGN FROM THE MENU. GUESSES DO NOT HAVE TO BE PREFACED WITH ANYTHING, BUT **CERTAIN** RESPONSES MUST BE PRECEDED BY THE WORD "CERTAIN." THIS WILL ALLOW ME TO KEY IN THE RESPONSE AS QUICKLY AS POSSIBLE FOR ACCURATE DATA COLLECTION.

Let us commence the practice trials and I will answer any questions you have as we proceed. I will also prompt you on your response technique, as a reminder.

APPENDIX C

TABULAR DATA TO ACCOMPANY FIGURES 55 THROUGH 58

TABLE 46

Summary of differences in mean legibility distance for subsets of data, young drivers only, field study data - Figure 55

Warning Sign	First Response*	Mean of First Two Responses*	Minimum of Six Responses*	Mean of Six Responses*
Two-Way Traffic (W6-3)	341.0 m (1118.7)	343.1 m (1125.5)	304.9 m (1000.4)	355.5 m (1166.3)
Bicycle Crossing (W11-1)	187.9 m (616.5)	201.2 m (660.1)	178.1 m (584.3)	253.1 m (830.4)
Crossroad (W2-1)	354.1 m (1161.8)	362.6 m (1189.6)	341.8 m (1121.3)	382.8 m (1255.8)
Deer Crossing (W11-3)	318.2 m (1044.0)	307.2 m (1008.0)	268.2 m (879.8)	335.2 m (1099.9)
Slippery When Wet (W8-5)	271.8 m (891.7)	278.6 m (913.9)	248.0 m (813.8)	320.6 m (1051.7)
T-Junction (W2-4)	350.9 m (1151.4)	359.2 m (1178.6)	325.8 m (1068.9)	405.5 m (1330.3)

* **numbers in parentheses denote feet**

TABLE 47

Summary of differences in mean legibility distance for subsets of data, older drivers only, field study data - Figure 56

Warning Sign	First Response*	Mean of First Two Responses*	Minimum of Six Responses*	Mean of Six Responses*
Two-Way Traffic (W6-3)	245.3 m (804.8)	244.8 m (803.3)	214.3 m (703.2)	261.1 m (856.5)
Bicycle Crossing (W11-1)	123.5 m (405.3)	128.2 m (420.5)	117.3 m (384.9)	156.9 m (514.8)
Crossroad (W2-1)	283.7 m (930.7)	287.3 m (942.5)	252.5 m (828.3)	289.6 m (950.0)
Deer Crossing (W11-3)	166.2 m (545.3)	169.9 m (557.3)	139.7 m (458.4)	205.1 m (673.0)
Slippery When Wet (W8-5)	139.8 m (458.8)	157.9 m (518.0)	125.0 m (410.0)	188.2 m (617.3)
T-Junction (W2-4)	271.5 m (890.6)	275.1 m (902.4)	246.8 m (809.8)	295.0 m (967.9)

* **numbers in parentheses denote feet**

TABLE 48

Summary of differences in mean legibility distance for subsets of data, young drivers only, laboratory study data - Figure 57

Warning Sign	First Response*	Mean of First Two Responses*	Minimum of Six Responses*	Mean of Six Responses*
Two-Way Traffic (W6-3)	381.5 m (1251.7)	371.0 m (1217.1)	326.9 m (1072.5)	369.8 m (1213.2)
Bicycle Crossing (W11-1)	266.5 m (874.2)	263.1 m (863.3)	211.3 m (693.3)	282.2 m (925.8)
Crossroad (W2-1)	429.2 m (1408.3)	415.4 m (1362.9)	390.9 m (1282.5)	422.9 m (1387.4)
Deer Crossing (W11-3)	356.9 m (1170.8)	337.8 m (1108.3)	303.5 m (995.8)	359.4 m (1179.0)
Slippery When Wet (W8-5)	342.9 (1125.0)	352.2 m (1155.4)	288.9 m (947.8)	364.8 m (1196.7)
T-Junction (W2-4)	434.9 m (1426.7)	436.6 m (1432.5)	392.2 m (1286.9)	434.7 m (1426.0)

* **numbers in parentheses denote feet**

TABLE 49

Summary of differences in mean legibility distance for subsets of data, older drivers only, laboratory study data - Figure 58

Warning Sign	First Response*	Mean of First Two Responses*	Minimum of Six Responses*	Mean of Six Responses*
Two-Way Traffic (W6-3)	242.1 m (794.2)	250.9 m (823.3)	209.0 m (685.8)	265.4 m (870.6)
Bicycle Crossing (W11-1)	133.4 m (437.5)	136.4 m (447.5)	116.6 m (382.5)	158.7 m (520.7)
Crossroad (W2-1)	305.0 m (1000.8)	311.7 m (1022.5)	280.9 m (921.7)	319.9 m (1049.7)
Deer Crossing (W11-3)	262.4 m (860.8)	254.1 m (833.8)	215.6 m (707.5)	258.3 (847.4)
Slippery When Wet (W8-5)	213.4 m (700.0)	226.3 m (742.5)	149.6 m (490.8)	251.8 (826.1)
T-Junction (W2-4)	302.5 m (992.5)	316.0 m (1036.7)	279.3 m (916.2)	327.4 (1074.0)

* **numbers in parentheses denote feet**