Final Report: Intersection and Pedestrian Safety Research BD543-17

July 2009

Neil Charness, Michael Champion, Ainsley Mitchum & Mark Fox

Department of Psychology, Florida State University

Final Technical Report No. BD543-17

Department of Psychology

Florida State University

Tallahassee, FL 32306-4301

Technical Report submitted to the Florida Department of Transportation, Tallahassee under contract BD543-17 to Neil Charness, Principal Investigator.

SI* (Modern Metric) Conversion Factors

Approximate Conversions to SI Units

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
LENGTH				
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
AREA				
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
VOLUME				
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
NOTE: volumes greater than 1000 L shall be shown in m^3				
MASS				
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
TEMPERATURE (exact degrees)				
^oF	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	^oC
ILLUMINATION				
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
FORCE and PRESSURE or STRESS				
lbf	pound force	4.45	newtons	N
lbf/in^2	pound force per square inch	6.89	kilopascals	kPa

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
LENGTH				
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
AREA				
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
VOLUME				
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
MASS				
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
TEMPERATURE (exact degrees)				
^oC	Celsius	1.8C+32	Fahrenheit	^oF
ILLUMINATION				
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
FORCE and PRESSURE or STRESS				
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380. Source: http://www.fhwa.dot.gov/aaa/metricp.htm (Revised March 2003)

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Intersection and Pedestrian Safety Research		5. Report Date July 2009
		6. Performing Organization Code
7. Author(s) Neil Charness, Michael Champion, Ainsley Mitchum, & Mark Fox		8. Performing Organization Report No.
9. Performing Organization Name and Address Florida State University, Tallahassee, Florida 32306-4301		10. Work Unit No. (TRAIS)
		11. Contract or Grant No. BD543-17
12. Sponsoring Agency Name and Address Florida Department of Transportation		13. Type of Report and Period Covered Final Report 2007-2009
		14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

Eight studies were conducted using lab and field experiments (day and night) to assess sign perception and pedestrian awareness for younger, middle-aged, and older drivers, and to assess older and younger pedestrian intersection crossing behavior. Study 1 showed that fluorescent yellow sign sheeting marginally enhanced perception and comprehension distance over standard sheeting. Age strongly predicted sign perception response time and perception distance. Lab-based measures predicted field performance well. Study 2 showed superiority of uppercase street sign names over mixed case, particularly for longer names and for older drivers. Lab to field prediction was weak. Study 3 showed an advantage for symbol over word warning signs, and weak prediction from lab to field performance. Study 4 showed that older drivers comprehended portable changeable sign messages at about 870 feet, below the MUTCD guideline of 0.5 miles visibility. Study 5 used film clips with animated countdown signals and novel pie signals. Older pedestrians did not differ from younger ones for crossing decisions, though exhibited less comfort. Pie displays promoted lower likelihoods for decisions to cross long intersections compared to short and intermediate ones; the opposite occurred for countdown displays. Study 6 showed adequate legibility (300 feet) for countdown displays in the field. Older pedestrians crossed less often and were more likely to fail to cross in time than younger ones. Study 7 showed that all drivers were sensitive to the presence of a pedestrian during left and right turns, indicated by vehicle velocity profiles, though older drivers drove more slowly. Study 8 showed much reduced sign perception distances with simulated rain conditions on the vehicle windshield, though no age differences in performance, and an advantage for symbol over word warning signs.

17. Key Word	18. Distribution Statement
Pedestrian, driver, older driver, signal, countdown signal, age, sign perception, fluorescent, portable changeable message sign, lab study, field study	No restrictions

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	100	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Disclaimer

Acknowledgments

(1) Prepared in cooperation with the State of Florida Department of Transportation and the U.S. Department of Transportation. We thank our Project Manager, Gail Holley, for her advice and patience, as well as for providing access to signs.

(2) We thank Blueprint 2000 for permission to use Broadmoor Estates, and Susan Newhouse at the Tallahassee Police Department for help in scheduling officers, as well as the many officers who served at that site.

(3) We thank David Rodrick, who started as a postdoctoral fellow on the project (now at University of Michigan, Dearborn) and the many Florida State University psychology undergraduates who worked on this project: Payton Huhta, Hillary Wolski, Cristina Del Busto, Stephan Pryzbylowicz, Mark Basista, Jonatan Alvernaz, Alan Lemura, Jeremy Jarrett, John Reis, Kristin Bruton, Darrow Speyer, Kelly Schietert, Kaitlyn Sisson, Marilyn Ramos-Deese, Michael Jones, and Tatyana Burgo. We are extremely grateful to all the drivers and pedestrians who participated in this research.

Executive Summary

Florida has one of the oldest state populations in the U. S. with over 3 million of its roughly 18 million citizens estimated to be age 65 or older as of July 2008, and over 520,000 estimated to be age 85 or older. Given the greater vulnerability of these citizens to crashes, both as drivers and as pedestrians, the state is trying to provide road environments that can reduce driver and pedestrian error and thereby maximize safety, in accord with the goals of the Florida Department of Transportation's older driver program, 'Safe Mobility for Life.' We made use of human factors techniques, including lab and field studies, to assess person-environment fit characteristics.

The perspective we have taken, particularly for road signs and traffic signals, is that an effective sign will be one that can attract attention, be legible, and be comprehensible soon enough that the observer can safely take appropriate action. In a set of 8 studies, we investigated how sign characteristics under day and night conditions and different weather conditions affected their comprehensibility as well as the decision making processes of drivers and pedestrians. A unique aspect to this project was the employment of a combination of lab-based, and field-based tasks to enable us to predict from the former to real world settings for younger, middle-aged, and older driver and pedestrian populations.

Study 1 examined how well warning signs were identified in both field and lab studies as a function of the type of sheeting and the age of the driver. We used both pattern matching (replica of sign shown to the driver) and comprehension probes (a verbal description is given to the driver). Critical variables were the age of the driver (young, middle-aged, older) and sheeting for the warning sign (fluorescent, standard). Consistent with earlier findings in the literature we showed a small, though non-significant advantage for fluorescent sheeting over standard sheeting. There was a hint that the effect might be partially mediated by intensity of headlamps, with fluorescent sheeting doing better under low beams and perhaps worse under high beams.

Our second goal for this set of studies was to determine if lab-based analogs of field-based tasks accurately predict field-based performance. We found that this is indeed the case for lab-based pattern matching (or comprehension) response time measures when predicting field-based stopping distance for warning signs. The lab-based pattern matching task seems to be the most sensitive one so we would advise the adoption of that procedure. The most popular field procedure is naming, rather than pattern matching or comprehension, so additional research may be needed to assess if the successful prediction holds for that field based task. However, the factor analysis of lab and field based measures give us some confidence that these procedures are tapping a common construct, warning sign perception.

In both studies there were robust effects of age group, with older drivers taking longer or moving closer to the warning signs before comprehending or matching them.

In study 2 we examined how well street signs were identified when viewed in laboratory and field settings by younger, middle-aged, and older drivers. We examined how the two factors of street name length and case affected sign legibility and comprehensibility. The signs varied in number of letters in the name (6, 8, 10 in the lab; 8, 10 in the field) and in case (uppercase, mixed case). For the field study, drivers viewed the signs in daylight and at night.

In both field and laboratory studies, accuracy remained relatively high across age groups for drivers making decisions about whether a street sign matched or did not match a target name as a function of the number of letters in the name and whether the

lettering was all uppercase or mixed case. However, driver age was an important predictor of how soon (distance from sign in the field, response time in the lab) drivers made their decision. Older drivers were slower than middle-aged ones who in turn were slower than younger drivers.

The distance from the sign (field) and the speed of that decision (lab) were also affected by sign characteristics. Drivers responded most quickly to six-letter signs, then 8-letter signs which in turn were faster than 10-letter signs, suggesting that number of letters in a street name be minimized. Uppercase signs were somewhat superior to mixed case signs and we would advise their adoption. In the field study, responses at night were considerably slower than responses during the day. The weak correlation between outcome measures across study types indicates that field studies are needed to validate the findings in lab studies.

In study 3 we evaluated the legibility of symbol versus word warning signs using comprehension and pattern matching procedures in a lab task and a comprehension procedure in a field task. For the lab studies, there were clear advantages to symbol signs at the simulated farthest distances (200, 500 feet) for both accuracy and speed of responding whether pattern matching or comprehension measures were used. In general, younger drivers responded more quickly than middle-aged and older drivers.

For the field study we investigated legibility of the same warning signs in day and night conditions using a comprehension measure. Signs were comprehended at greater distances (by about 80 feet) during the day than at night. Symbol signs were on average much more easily comprehended than signs with word messages, holding an advantage of about 300 feet in legibility distance. Age was not a significant factor in legibility distance, though our relatively small sample size in this study did not afford much power to detect age differences or interactions of age with other variables.

There was weak prediction from lab study performance to field study performance, indicating the need for field research to provide policy guidelines. Nonetheless, based on both lab and field results, we recommend the use of symbol warning signs rather than warning signs with word messages.

Study 4 assessed whether portable changeable message signs (PCMS) were comprehended equally well at 900 feet by older and younger drivers in day time viewing conditions. We used a comprehension procedure, asking drivers to determine if a description of the sign message matched the message on the PCMS (same / different judgment). They performed this task on foot, walking toward the sign until they could make their decision.

Accuracy was relatively high in this judgment with older adults slightly, though non-significantly, worse. However, there was a marked age difference in the distance at which the judgment took place, with older drivers walking closer to the sign by about 30 feet and younger drivers making their judgments from the full 900 feet. Thus, older drivers in particular were not able to understand the messages at the Manual on Uniform Traffic Devices (MUTCD, 2003) recommended distance of 0.5 miles. These results indicate the need to find ways to improve legibility of messages for PCMS.

Study 5 assessed younger and older pedestrians' crossing decisions for, and comfort with 3-second animated clips of short, medium, and long intersections displaying either a pedestrian countdown timer with flashing hand display or a novel animated pie display as a control condition. Older pedestrians were significantly less comfortable with crossing decisions though they made choices equivalent to those of younger pedestrians. The countdown display yielded fewer positive crossing decisions on shorter than on intermediate and longer intersections. The pie display generated slightly more positive decisions though more so for short and intermediate intersections and fewer for long intersections. We recommend that further research, particularly field studies, be conducted into the factors influencing pedestrian crossing decisions.

Study 6 assessed legibility of countdown signals for older and younger pedestrians. It also examined decision-making about entering short (50 ft), medium (75 ft), and long (115 ft) intersections with full or partial time displayed on a countdown signal. Legibility accuracy was uniformly high across age groups (94%) with older pedestrians making their decision about sign state at a significantly shorter distance (307 ft versus 375 ft) than younger ones. However, standard countdown signals are highly legible at distances greater than most intersections for both younger and older pedestrians so current guidelines for their construction are adequate for these age groups.

Older pedestrians, despite being more conservative about initiating a traversal of an intersection than younger ones were judged more likely to fail to cross in time than younger pedestrians for longer and medium-length intersections. We recommend that conservative timing intervals, assuming a slower walking speed than 4 ft/s be adopted for intersections frequented by older pedestrians to ensure that they have adequate time to cross safely.

Study 7 assessed whether younger and older drivers are equally sensitive to a pedestrian encroaching on an intersection in daylight. Drivers were asked to make left and right turns with a pedestrian present and approaching the road from the lane opposite that used by the driver during the turn, or not present near the intersection. We used velocity profiles from radar gun data to assess the sensitivity of the driver to the type of turn and the presence of the pedestrian.

Drivers of all ages showed some sensitivity to the presence of a pedestrian entering the lane opposite that of the driver at an intersection. However, there was little difference between older and younger drivers in response to the presence of a pedestrian. Older drivers drove more slowly on the track than did younger ones, and showed greater sensitivity to whether the turn they were to make was a right or left turn. Thus, for the case of a pedestrian wearing normal street clothes entering an intersection there is little difference in the behavior of a younger or older driver. Older drivers tend to drive more slowly on average, particularly for right turns.

Given that the literature indicates that many pedestrian crashes occur at night, we recommend further field studies examining how driver velocity profiles change in the presence of pedestrians varying in conspicuity.

In study 8, we simulated heavy rainfall on a vehicle windshield and assessed sign legibility as a function of sign type (symbol, word) and driver age (young, older) using a sign comprehension measure. Age was not a significant factor in the comprehension of signs under adverse weather conditions (rain simulation). Symbol signs maintained a large (60 foot) legibility distance advantage compared to signs with word messages. Thus, we strongly recommend the use of symbol warning signs rather than ones with word messages to improve legibility under adverse weather conditions.

Table of Contents

List of Tables

List of Figures

.

Chapter 1.

Introduction

The US population is aging and Florida has one of the oldest state populations. Over three million of its roughly 18 million citizens (17.4%) were age 65+ years as of July 2008, and over 520,000 were age 85+ years (http://www.census.gov/popest/states/asrh/SC-EST2008-01.html). In some Florida counties (e.g., Sarasota, Charlotte) 30 percent were age 65+ in 2007 (http://quickfacts.census.gov/qfd/states/12/12103.html). With the aging of the baby boom cohorts, we face a growing challenge to design safer road and street environments for Florida's citizens, and for those in other states (Transportation in an aging society, 2004). Why focus on older adults? For one, older adults are more vulnerable to the effects of crashes than younger adults. For the same intensity crash, 80 year old drivers are~ 4-5 times more likely to die in the collision than 20 year old ones (Evans, 2004). Similarly, even though older adults represent a relatively small fraction of the general population, they are likely overrepresented in pedestrian deaths because of their greater fragility, at ~4 times the fatality rate per 100,000 people (Dunbar et al., 2004; Evans, 2004; Oxley, Fildes, & Dewar, 2004). However, unlike the case for motorist fatalities where risk can be quantified based on national surveys of driver mileage, how best to assess pedestrian risk is unresolved (e.g., Emo & Do, 2008).

Approaches to improve safety include better screening of older drivers for fitness to drive (though this approach will not work for older pedestrians), better training of older drivers to improve their driving skills, and better design of pedestrian and driving environments. Current instruments to predict crashes in older drivers are simply not precise enough (Bédard, Weaver, Darzins & Porter, 2008). Current popular courses for older drivers are not effective in reducing crash rates (Nasvadi & Vavrik, 2007). Thus, improving road environments to prevent and reduce driver and pedestrian error seems warranted.

Although to err is human (Kohn, Corrigan & Donaldson, 2000; Sharit, 2006), better design of existing transportation systems can contribute to safety. Also, designs that foster safe conditions for older adults tend to aid children too. As an example, intersection crossing time is set on the assumption that people walk at a speed of about 1.2 meters/second, but women over the age of 80 typically walk at a rate of about 1.1 m/s (Steenbekkers & van Beijsterveldt, 1998). Allowing additional time at a traffic intersection for older adults to cross safely will also aid small school children who have similar slow walking rates, though at the cost of slowing motorists (Charness & Bosman, 1990). Also, older adults are a promising model for both international tourists to Florida and our non-English speaking population. Both those groups can be somewhat compromised in speed of reading and comprehending of signs. Enhancing the safety of roadways involves a cost and determining whether the costs are commensurate with social welfare benefits is always a difficult policy decision. Our aim is to provide relevant data to inform decision making.

The perspective we have taken, particularly for road signs and traffic signals, is that an effective sign will be one that can attract attention, be legible, and be comprehensible soon enough that the observer can safely take appropriate action (e.g., Fisher, 1992). This corresponds approximately to the Manual of Uniform Traffic Control Devices (MUTCD) guidelines about time for processing warning signs: the Perception, Identification, and Emotion stages (from the PIEV time, also now called perception-response time). That is, both physical characteristics of the sign (legibility) and perceptual and cognitive capabilities of the perceiver (comprehension and decision factors) affect driver and pedestrian behavior. Usually, the relevant actions to be taken involve decisions about braking, stopping, accelerating, and lane changing for drivers, and the decision to enter and navigate an intersection for pedestrians. Critical factors for legibility of signs and signals include size, foreground/background contrast, font, color, ambient light/weather conditions, and position of the observer.

A critical individual difference factor affecting speed of comprehension and decision making in these studies is the age of the observer. Aging degrades perceptual, cognitive and psychomotor functioning (Fisk et al., 2009). There is ample evidence that these changes affect perception of traffic signs in laboratory experiments (Dewar, Kline, Scheiber & Swanson, 1997) where legibility distances for older drivers were typically about 80% of those for younger ones. Thus we sample younger, middle-aged, and older adults in most studies and compare younger versus older adults in others. There are a number of methodologies that have been used to assess sign perception and comprehension, ranging from highly constrained laboratory studies to simulator studies to field studies. Each methodology has strengths and weaknesses. Laboratory studies have high internal validity given tight control over experimental variables, but have questionable external validity (ability to generalize to real world settings). Field studies have high external validity but may lack internal validity. Simulator studies represent a case midway between lab and field, though problems with simulator sickness can prevent many older adults (up to 40%) from participating in such studies (Edwards et al., 2002). We use both lab and field studies.

Objectives and Supporting Tasks

Both pedestrians and drivers require proper signage in order to navigate safely. The purpose of this project has been to improve our knowledge about sign legibility in order to develop appropriate guidelines for the Florida Department of Transportation (Department) and for local governments that depend on their recommendations. This research program relies on human factors principles and methodologies. These methodologies are critical to being able to fulfill the initiatives outlined in the Safe Mobility for Life Program business plan:
http://www.dot.state.fl.us/trafficoperations/Operations/PDFs/Safe_Mobility_Booklet%20final.pdf).

We followed three interlocking projects, with an emphasis on meeting the needs of younger, middle-aged, and older Floridians:

(1) Functionality of traffic signs and signals for motorists of varying ages.
(2) Functionality of traffic signals for pedestrians of varying ages.
(3) Guidelines for signage to accommodate ambient light and weather extremes.

A unique aspect to this project was the combination of lab-based tasks and field-based tasks to enable us to predict from the former to real world settings for a range of younger, middle-aged, and older driver and pedestrian populations. Most projects were carried out at the Traffic Engineering and Research Lab (TERL) 2612 Springhill Road, Tallahassee, Florida, supplemented by lab-based tasks at Florida State University. We found it necessary to use another facility, Broadmoor Estates, an appropriated housing neighborhood[1], to provide adequate distance to warning signs for several studies.

[1] We thank Blueprint 2000 http://www.blueprint2000.org/about.html for permission to use Broadmoor Estates.

Chapter 2.

Study 1: Perception & Comprehension of Fluorescent Yellow Warning Signs

The Transportation Research Board has recommended the use of microprismatic fluorescent sheeting to improve the legibility of warning signs (NCHRP Report 600A: Human Factors Guidelines for Road Systems http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rpt_600B.pdf). There are several studies that support the use of fluorescent sheeting on warning signs to improve their visibility. However, initial studies used observers outside of vehicles to make judgments about signs or had them sit in the passenger seat of a test vehicle, or sit in a railway car (Jenssen et al., 1996). For instance, Schnell, Bentley, Hayes & Rick (TRB Paper 01-2417) noted a relatively small advantage for fluorescent sheeting of between 5 and 16% under good light conditions (daylight, during a fair day), but commented that this advantage was slight and might increase under less favorable viewing conditions. The problem, from a policy perspective, is whether the increased cost of this sign sheeting compared to standard sheeting is worthwhile, that is affects collision data.

Eccles & Hummer (TRB Paper No. 01-2236) have shown a slight advantage at 4/7 sites using a before/after study where standard signs were replaced by fluorescent ones. However, such studies are difficult to evaluate given that they are observational, and changes in traffic patterns (increasing/decreasing traffic density) might account for changes in observed collision variables. One area where there may be an important advantage for fluorescent sheeting is with older drivers. A study by Anders (2000) involved younger and middle-aged to older (age 55+) drivers, in daylight and night time conditions with a number of different combinations of signs (using the phrase “test route” with a forward arrow), with one being non-fluorescent yellow on non-fluorescent purple. Motorists drove by signs on a highway in an experimental test vehicle (using low beams at night). Analyses on late braking behavior and wrong turns (taken as an index of sign perception difficulty) showed no effects of sign type but one problem with this study is that there was only a single sign tested once. However, driver preference ratings showed a strong preference for the fluorescent signs with the higher contrast ratios (character to background ratio). Older drivers did not differ much from younger ones except for rating signs in general easier to perceive.

Chrysler, Stackhouse, Tranchida & Arthur (2001) had older drivers (mean age 71) read experimental street signs in early evening hours in winter in St. Paul, MN at two different intersections in an instrumented car and estimated legibility distances for Type IX, Type VII (microprismatic sheeting), Type III (encapsulated lens material) and Type I (lens retroreflective) sheetings. They found greater legibility distances for the retroreflective Type VII (170 feet) and Type IX (172 feet) over Type III (142 feet) which was superior to Type I. Interactions with type of intersection and placement (right, left sides of roadways) indicated a larger advantage for the microprismatic signs.

Chrysler, Carlson & Hawkins (2002) tested drivers aged 55-75 at night only on a closed course using a test vehicle. They tested signs with word messages (e.g., containing text such as: area, city, deer, road, when, etc.). They tested Type III, Type VII, and Type IX signs varying background colors (green, orange, white, yellow) and foreground colors (black, white). Drivers were asked to say the word on the sign. Again, they found an advantage for the microprismatic signs (184, 183 feet compared to 174 feet for Type III). They also found advantages for yellow and white signs. They also found an unexpected advantage for Standard Highway Series D fonts over Clearview ones, though as they pointed out signs were in all uppercase, varying from prior studies that used mixed case. We address this case issue in one of our studies.

Charlton (2006) has argued that conspicuity (whether a sign is noticed) and static comprehension are the two most important features of hazard warning signs. Hence, of the many possible variables that can be used to indicate effects of signs, we have chosen to focus on those aspects.

Our study aimed to improve on some of the deficiencies in prior ones by letting drivers approach signs in their own vehicles, a situation that should generalize better to typical driving conditions. We presented multiple trials with different signs to get better estimates of accuracy. We also tested younger, middle-aged and older adults to assess relative advantages and disadvantages in a broad age range. We also attempted to look at the effect of using different instructional paradigms, both basic sign perception (sign matching) and sign comprehension processes, in part because Zwahlen and Schnell (1999) have indicated that the type of instructions may be critical to interpreting experimental results. Drivers were asked to approach signs until they could answer a relevant question about the sign. Half the time the answer was intended to be yes (match) and half no (mismatch).

One of the important issues to resolve to understand the legibility of street signs is the length of the sign given that time to read is going to be somewhat proportional to the length of the message. In a preliminary survey of Leon County street signs we observed the following (Figure 1) distribution of sign lengths.

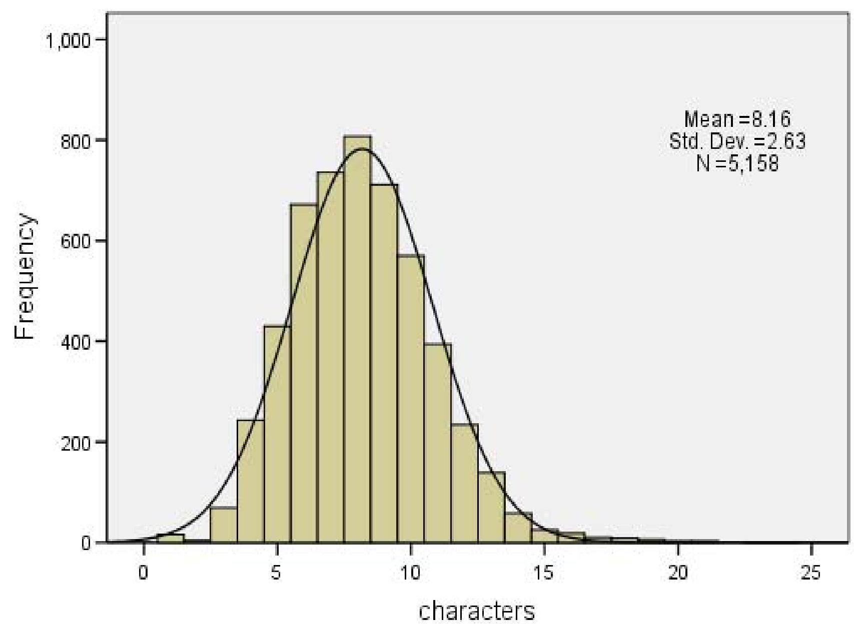


Figure 1. Histogram of Leon County street name lengths (number of characters)

The mean length was 8 letters, and the distribution is approximately normal, though somewhat skewed right. Most signs fall between the 25th and 75th percentiles, corresponding to lengths of 6 and 10 letters. In our studies, we chose name lengths of 6, 8, and 10 letters accordingly.

As mentioned above, there is controversy over whether laboratory studies can generalize well to field studies. We addressed this issue by pursuing roughly parallel laboratory and field studies. In the laboratory studies we concentrated on an outcome measure -- response time -- that is difficult to measure without instrumented vehicles in field studies. Response time should give a fairly strong indication of the *minimal* time needed to process a sign or signal. Thus, any differences observed for variables such as sign type (e.g., symbol versus word), sign size (a surrogate of distance from sign in the field), sign case (mixed case versus uppercase lettering), and individual difference variables such as age should be valid indicators of *minimal/optimal processing time* in a transportation environment. Realistic response times when people are not pre-attending to sign information would be expected to be longer, a reason why more extreme values, usually 85^{th} or 15^{th} percentiles in the population, are often used to set standards. A value of 2.5 s is often adopted as the average perception-response time for motorists to take action from a sign (Manual on Traffic Control Devices, 2003)

For field studies, we concentrated on legibility and comprehension distance: the minimal distance from the sign or signal that was necessary to be able to respond accurately to a probe that asked the driver to match a target that was a replica of the sign (matching condition), or that was a comprehension probe (had the same meaning as the target). The legibility/comprehension distance variable is expected to provide a *best case* indicator of the ability to process the message that a sign conveys. It is best case because in more visually complicated environments, ones where the driver is processing all the features of the driving environment including other vehicles, pedestrians, and commercial and traffic signs, they are less likely to notice and to be actively processing sign information. Realistic legibility/comprehension distances would be expected to be longer. However, the picture conveyed from the lab and the field studies should provide a more thorough examination of sign/signal processing than either in isolation. It is also possible to look at the correlation between these two methodologies to see to what extent one might substitute for the other. Laboratory studies are much less costly to conduct than field studies.

Method.

Participant screening

For all participants, the requirements for inclusion were that they had a valid driver's license, owned a vehicle, were able to drive at night, drove at least twice a week, and did not show any significant deficits in intellectual functioning. All participants completed the following screening measures by telephone to determine eligibility for inclusion: The Short Portable Mental Status Questionnaire (SPMSQ) and Logical Memory I (LM I) subtest from the Wechsler Memory Scale – Revised Edition (WMS-R). In order to be eligible to participate, those with up to 12 years of education could make no more than 2 errors on the SPMSQ, while those with 13 years of education or more could make no more than 1 error. On the LM subtests from the WMS-R participants were required to earn a minimum of 7 points out of 25 on story 1 in order to be included. If participants failed story 1, story 2 was administered. If the participant then earned a score of 7 points out of 25 points on story 2, they were eligible to be included in the study. These tests were meant to screen out those with severe memory problems or dementia. Participants screened into the study filled out an IRB-approved informed consent document when appearing in person for the experiment or field study and also a questionnaire providing demographic information.

Study 1 Laboratory Task.

Design

The laboratory study was designed to examine correlations across laboratory and field using similar tasks. A 3 age (younger, middle, older) x 3 distance (20 feet, 200 feet, 500 feet) x 2 task (pattern matching, comprehension) mixed factorial design was used with age as a between-subjects variable, and distance and task as within subject variables. Dependent variables were reaction time and accuracy.

Participants

Twenty younger (M = 24 yr), 19 middle-aged (M = 58 yr), and 20 older drivers (M = 75 yr) participated in this study and also in the field study. Although gender was not an analyzed variable due to the small sample size, we attempted to balance male and female participation (10 per group). Participants were recruited, screened, and paid in the same manner described in Study 1 field.

Materials and Procedure

Participants were tested individually and completed five tasks in the following order except when otherwise specified: a demographic questionnaire, pattern matching and comprehension computer simulations of Study 1 Field, a visual acuity test, and a digit symbol-coding task. The demographic questionnaire and digit symbol-coding task were completed with paper and pencil, and the simulation tasks and visual acuity test were administered on computer. Pattern matching and comprehension were counterbalanced for order, but always occurred second and third in the sequence of tasks. The entire procedure lasted approximately 45-60 minutes. This, and all other lab experiments (with the exception of Experiment 5) were conducted via E-Prime software (http://www.pstnet.com/).

Demographic Questionnaire

The fourteen-page questionnaire solicited information on demographics, general health, vision health, driving habits, and depression symptoms.

Pattern Matching

Participants were seated and situated with their eyes approximately 34 inches from a 19-inch, 60 Hz., CRT monitor. Each trial began with a one-second presentation of an image of one of four signs from Study 1 Field (stop ahead [W3-1], yield ahead [W3-2], speed zone ahead [W3-5], signal ahead [W3-3]) in the middle of the screen. Signs in this initial presentation occupied about ten degrees of visual angle and corresponded to a 20ft. viewing distance. A 500ms blank screen (white) immediately followed the disappearance of the sign before a second image of a sign appeared on the right-hand side of the screen at either the top, middle, or lower portion of the screen. This second sign appeared in one of three sizes, corresponding to distances of 20, 200, and 500 feet (latter two were more precisely 198 and 502 feet based on visual angle estimates for normal viewing distance from the screen). Participants pressed keys labeled “same” or “different” after judging if this sign and the initial sign were similar or different. Participants were instructed to make this forced choice as quickly as possible while sustaining accuracy.

All 72 trials occurred in a fixed pseudo-random order with each initial sign displayed 18 times, and followed by an equal number of matching and mismatching signs. Pilot testing was conducted to pair each sign with a single mismatching sign that elicited the longest possible response times on mismatch trials. The pairings were stop ahead and yield ahead; and speed limit ahead and traffic light ahead. Prior to the task, participants practiced on four trials and received feedback of their accuracy. Feedback was not provided thereafter.

Comprehension

Comprehension was identical to pattern matching except for the following: Instead of an image of a sign, each trial began with one of four written statements that correspond semantically to the signs. These statements were presented in the middle of the screen for one second. Each statement corresponded to one sign. The statements were:

Stop ahead: Prepare to stop
Yield ahead: prepare to yield
Speed Zone ahead: Speed limit is reduced
Traffic light ahead: Traffic light ahead

The statements were printed in Courier 18-point font with each letter occupying roughly 0.5 degrees of visual angle. As in pattern matching, a sign appeared following a blank screen lasting 500 ms. Once again, the goal was to press “same” if the statement matched the meaning of the sign, and “different” if the statement and sign had different meanings.

Visual Acuity

Participants were given a short break if necessary before visual acuity was measured. An interactive visual acuity chart, consisting of the standard array of letters was displayed on a plasma monitor. Participants attempted to read the letters at a distance of ten feet. The test measures between 20/20 and 20/400 vision.

Digit Symbol-Coding

This task is administered with pencil and a single page of paper and measures processing speed. The task entails converting a long string of digits into symbols using a conversion key at the top of the page. There are seven rows of 20 digits with a blank box underneath every digit. Each digit from 1 to 9 corresponds to an abstract symbol that must be inscribed in the box underneath. The digits must be coded from left to right while advancing down the page line by line. The score is the total number of correctly transposed symbols after 90 seconds.

Signs

The signs used in this study are as seen in Figure 2.

W3-3	W3-5	W3-1	W3-2	W4-2R

Figure 2. Signs used within Task 1 laboratory and field tasks.

Results

Comprehension accuracy

Analysis of Variance (ANOVA) showed that there were significant ($p < .01$) effects of age group and distance as well as an age by distance interaction: $F(4, 112) = 6.07$, $p < .01$. In general, accuracy is high, >90%, though, as seen in Figure 3, the interaction illustrates that the effect of distance depends on age group, with young drivers showing minimal effects, middle-aged drivers showing moderate decline in accuracy with distance, and older drivers showing a strong decline in accuracy with distance, with moderate to high accuracy for 20 and 200 foot distances, but only 76% accuracy at 500 feet. Thus, our manipulation of sign size / distance was reasonably informative for the longer distances.

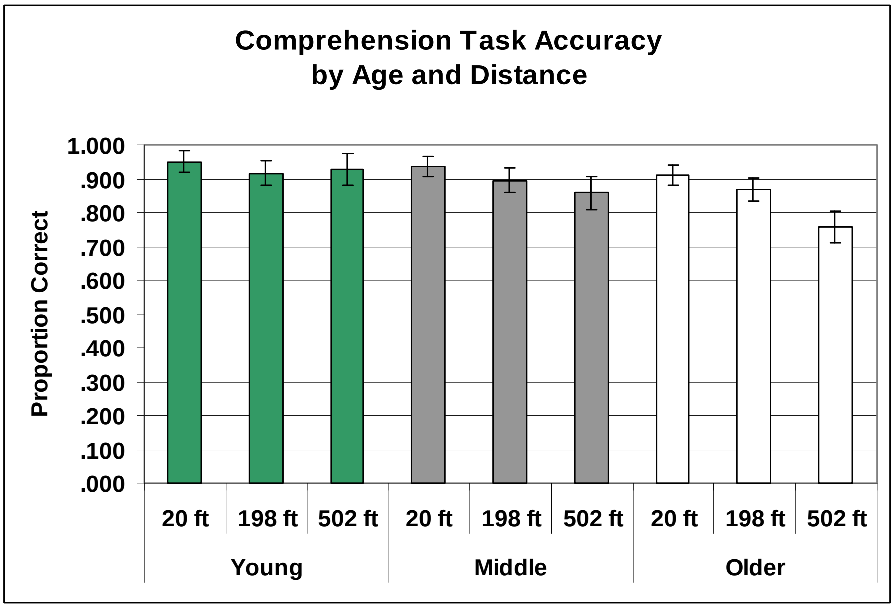


Figure 3. Comprehension accuracy (mean proportion correct) as a function of age group and simulated sign distance (size of sign). Error bars represent 95% confidence intervals for the means.

Given age-related enhanced sensitivity to distance (size) effects, response times will need to be interpreted carefully.

Comprehension response time

ANOVA showed significant effects of distance and age group, and a significant interaction, $F(4, 112) = 4.02$, $p < .01$, complementary to the results for accuracy.

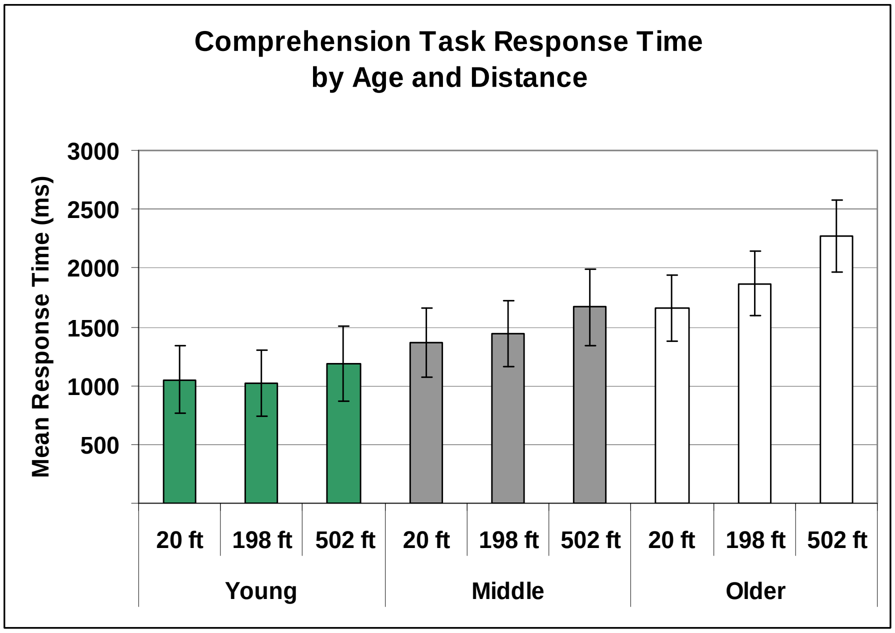


Figure 4. Comprehension response time (milliseconds) as a function of age group and simulated sign distance (size of sign). Error bars represent 95% confidence intervals for the means.

A seen in Figure 4, again, younger drivers show relatively little effect of distance on response time to a sign in the comprehension task (until it is at 502 feet), whereas middle-aged drivers show a moderate increase in response time with sign distance, and older drivers show a strong increase for the most difficult condition (sign size simulating a distance of 502 feet). At that distance, older drivers take about twice as long as younger drivers to make their decision about whether the sign matches the description of the target sign. Many guidelines for sign placement (e.g., MUTCD, 2007 revision) rely on a perception-response time (also called PIEV time) of 2.5 s. Note that older drivers begin to approach the 2.5 s cutoff (95% confidence interval for the mean includes that value) for just the perception component of sign comprehension. Deciding on and taking an action more complex than pressing the response key in the lab would add additional time.

Thus, both speed and accuracy measures follow similar trends, with the speed measure being somewhat more sensitive to the manipulated variables.

Pattern matching accuracy

Again, ANOVA showed significant main effects of distance as well as age group ($p < .01$) and a significant age group by distance interaction, $F(4, 112) = 5.80$, $p < .01$. As seen in Figure 5, the interaction is similar to those for the comprehension task, with little sensitivity to sign distance for younger and middle-aged drivers, but older drivers drop to 80% accuracy for 502 foot signs. In general, pattern matching is a more sensitive indicator variable than comprehension, in line with the lessened demands for interpreting the sign information.

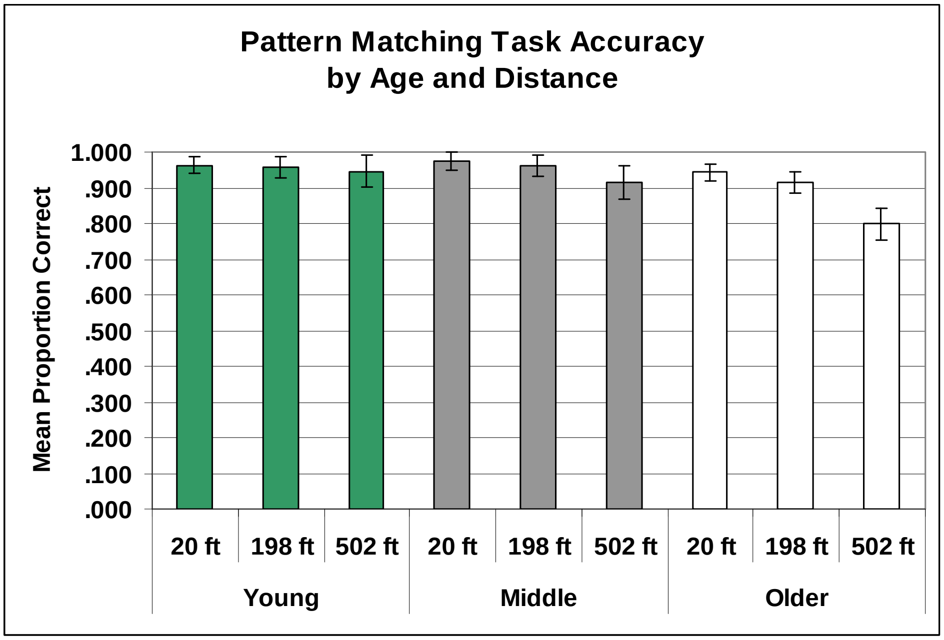


Figure 5. Pattern matching accuracy (proportion correct) as a function of age group and simulated sign distance (size of sign). Error bars represent 95% confidence intervals for the means.

Pattern matching response time

ANOVA shows the same pattern as for accuracy, with significant main effects ($p < .01$) of age group and sign distance, with a significant interaction of age and distance. As Figure 6 illustrates, the interaction shows that in younger drivers, the size of the distance effect is modest, whereas for middle aged drivers and older drivers it is pronounced for the most distant sign (502 feet). Note that response times are much quicker for a matching task than for a comprehension task, and the times even for older drivers are below 1.5 s.

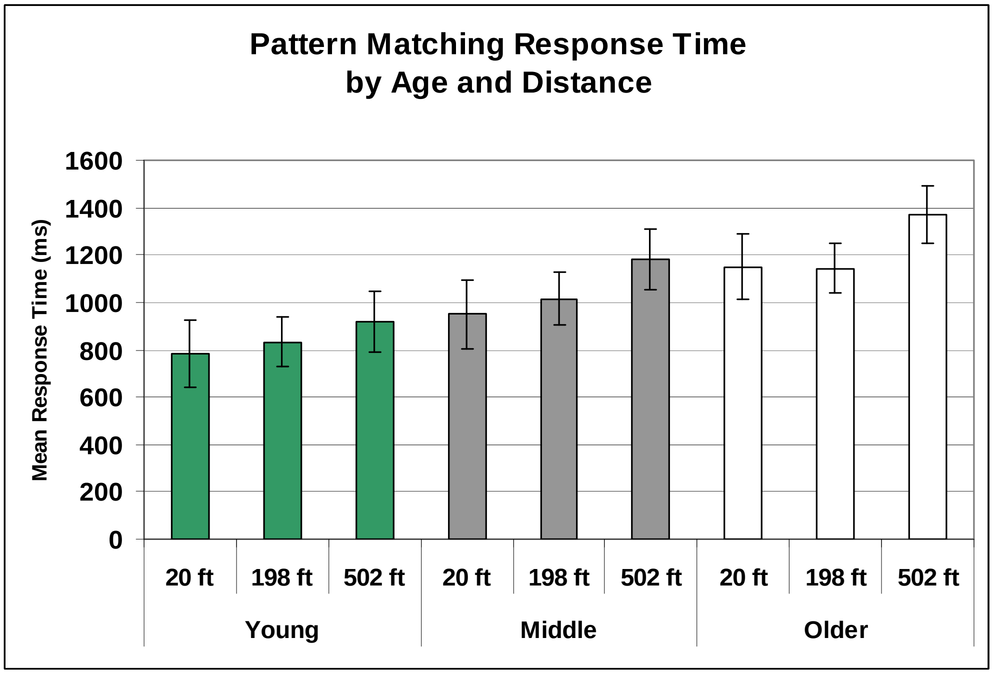


Figure 6. Pattern matching response time (milliseconds) as a function of age group and simulated sign distance (size of sign). Error bars represent 95% confidence intervals for the means.

Study 1 Field Task.

Design

The study used a 3 x 2 x 2 x 2 mixed design. Age (younger, middle aged, older) was the between subjects factor. Sign type (fluorescent vs. standard), condition (day vs. night), and task (pattern matching vs. comprehension) were within subjects factors. The order in which participants completed the pattern matching and comprehension tasks was counterbalanced between subjects. If a participant completed pattern matching first and then comprehension during their first visit to the test site, this order was reversed for their subsequent visit.

Participants

The participants were community dwelling older, middle aged, and younger drivers recruited from the Tallahassee, Florida area via newspaper ads and word of mouth. Six younger (M = 25 yr), 17 middle-aged (M = 58 yr) and 8 older (M = 71 yr) drivers participated in this study after completing the lab version. Participants were paid $10 per hour for taking part in the study.

Materials

Track

A closed track was used for all testing at Broadmoor Estates, an abandoned neighborhood scheduled for redevelopment by Blueprint 2000, Tallahassee. Testing was carried out with an off-duty police officer present to ensure safety of the drivers and experimenters. The area around the track was moderately shaded, similar to driving conditions on a rural road. See Appendix A. Any tree limbs or brush that directly obstructed visibility on the track were cleared before testing began. The track was level and straight, with a starting line marked at a distance of 1950 feet. Track markings indicating distance from the starting line occurred every 25 feet.

Signs

The stimuli consisted of a set of 10, 36 inch by 36 inch symbolic warning signs. The signs included are as follows (see Figure 1 as well): speed zone ahead (W3-5), stop ahead (W3-1), yield ahead (W3-2), right lane ends (W4-2R), and signal ahead (W3-3). Each sign occurred twice in the test set, once as a standard sign and once as a fluorescent, high intensity reflective sign. All signs followed MUTCD guidelines with the 5 standard signs using grade 7 retroreflective sheeting, while the 5 fluorescent signs used grade 9 diamond retroreflective sheeting.

Chroma-meter

Color space and luminance information was gathered using a Konica Minolta Chroma Meter CS – 100A. Information from this device was recorded.

Luminance meter

Luminance information not requiring color space coordinates were gathered from a Konica Minolta Luminance Meter LS – 100.

Radar gun

Track measurements were obtained using an Applied Scientific Stalker ATS radar gun. This device was placed behind the participant's vehicle and used time and speed to equate the location of the participant during each trial. Information was reported and recorded by a Gateway M – 285 – E laptop computer, and processed using Stalker ATS software.

Laptop computer

Two Gateway M285-E Laptop Computers were used in this experiment. Both computers are identical with Intel Centrino Core 2 Duo 1.6 GHz processors, 1 GB of RAM memory, and 80 GB hard drive.

Sign display

Signs were displayed on a custom designed sign changing apparatus (see Figure 7). The device stood a total of 12 feet tall. The bottom of the device was large PVC piping with rope wound around for stability and easy handling. The extending pole was eleven feet high with a mounting bracket for each tri- and duo-sign head. Sign mounting heads were affixed to all signs used in this study and were placed onto the device when required. The design of the mounting heads allowed 2 (square mounting heads) or 3 (triangular mounting heads) signs to be mounted at once. Once a set of signs was affixed, they could be rotated to present a different sign.

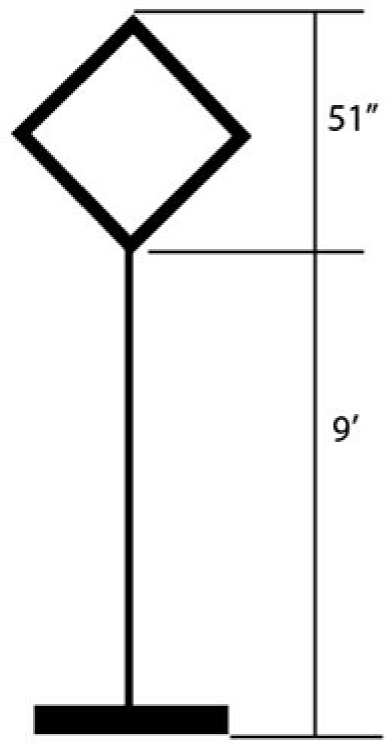


Figure 7. Diagram of Task 1 and Task 3 sign changing device.

Procedures

Day and night conditions

Participants were run through both day and night conditions, with some participating in the day condition first and others participating in the night condition first. The order in which participants were run through the day and night conditions was not varied systematically, but rather was dictated by scheduling constraints. The procedure for the day and night conditions were identical, with the exception of measurements that were taken before the experimental tasks began. These measurements will be explained in greater detail in the next section.

Measurements

Day condition

When participants arrived at the test site, the following measures were taken: Vehicle length, vehicle height, front wheel base to bumper, front bumper to participant, back bumper to participant, passenger side of vehicle to participant (Appendix A), participant height, and participant height to eyes (Appendix A).

Chromametric measurements were also taken both through the windshield and in open air to determine lighting conditions and the amount of light absorbed by the windshield. In addition, the make, model, and year of the participant's vehicle were also noted. Chromaticity readings were taken for each color on the signs at the beginning of each run. Sign colors included the following: Red, green, yellow, white, and black, with separate sets of readings taken for standard and fluorescent signs.

Measurements

Night condition

Although drivers were permitted to set their headlamps to their usual night time setting during the experimental trials (which could be either low or high beams), chromametric measurements were taken on both low and high beams. Headlight measurements were taken on a black and white board at a distance of 3 feet. Chromaticity readings were not taken on signs because the chromameter being used was not sensitive enough for readings in which light was less than 0.01 candela/meter^2.

Pattern matching

The pattern matching task consisted of 10 trials. For all trials, participants began at the marked starting line and were instructed to pull up until their front tires touched the speed bump at the starting line. At the beginning of each trial, participants were shown a picture of a warning sign on a laptop screen, then were instructed to drive forward at a comfortable speed until they could decide with 100% confidence whether or not the actual sign at the end of the track matched the sign they had been shown at the beginning of the trial. Participants were instructed to stop their vehicle once they were able to make this determination and await further instructions from the experimenter. For half of the trials, the picture matched the actual sign, and for the remaining trials, the picture did not match the actual sign.

This procedure differed slightly for the first 5 participants. For those participants they were shown a printed picture of the sign on an 8.5 x 11 sheet of standard white paper. This method of displaying stimuli was abandoned due to problems with low lighting in the night condition. For the comprehension task, the procedure remained the same throughout the study.

Comprehension

The comprehension task followed the same basic procedure as the pattern matching task. However, instead of being shown a picture of a road sign, participants were read a phrase indicating the meaning of a sign (e.g., "does the sign indicate stop ahead", "does the sign indicated yield ahead"), then asked to drive forward at a comfortable speed until they could determine with 100% confidence whether or not the actual sign at the end of the track had the same meaning as the phrase they had been read. For half of the total trials, the phrase matched the actual sign, and on other half, the phrase did not match the sign.

Distance and speed on both tasks was measured on each trial with a radar gun, with data recorded on a laptop computer. In the event of a radar gun fault, travel distance was determined by measuring from the nearest track marking to the rear bumper of the participant's vehicle. Additionally, on two randomly selected trials (1 fluorescent trial, 1 standard trial) during both the pattern matching and comprehension tasks, a measurement of sign reflectivity at viewing distance was taken. For these measurements, both the trial number and sign were recorded. On the final trial of both conditions, a track measurement was taken in order to confirm that the radar gun was accurately recording driving distances.

Results.

Participant attrition

We began with a lab study and given some delays in producing sign-changing equipment, the field study was not available for scheduling for several months. Thus, there was significant attrition across the lab and field studies with many participants dropping out after the lab portion of the study, and some not continuing in the field study from the day condition to the night condition. In subsequent studies, we attempted to schedule people with minimal delays through lab and field versions, and in night and day sessions. Better scheduling resulted in far less attrition.

Daytime Comprehension Accuracy

A total of 6 young (M = 25 yr), 17 middle-aged (M = 58 yr), and 8 older (M = 71 yr) drivers had complete data in the day task. ANOVA showed no significant effects of age, sign type (fluorescent, standard), nor their interaction, though trends were seen ($p < .10$) for age group, $F(2, 29) = 2.87$, and sign type by age group, $F(2, 29) = 2.53$. The direction of those trends was for older drivers to do better on fluorescent than standard signs (98%, 84%) with other age groups equivalent on the two sign types. Given the lack of significant effects (mean accuracy = 96%), stopping distance measures should be interpretable.

Night time Comprehension Accuracy

A total of 5 young, 14 middle, and 7 older drivers had complete data. With so low a sample size, we found no significant effects with ANOVA, nor any trends. Mean accuracy was 95%.

Daytime Pattern Matching Accuracy

We had complete data for 6 young, 17 middle-aged, and 9 older drivers. ANOVA showed no significant effects, with a mean accuracy of 96%.

Night time Pattern Matching Accuracy

We had complete data on 6 younger, 7 middle-aged, and 5 older drivers. We observed no significant effects with ANOVA, though there was a marginal ($p < .07$) age effect, with middle-aged drivers at 99% accuracy, and older and younger drivers at 90% and 93% accuracy respectively.

Day time comprehension stopping distance

We had complete data on 6 younger, 17 middle-aged, and 8 older drivers. ANOVA revealed that the only significant effect was due to age, $F(2, 28) = 4.11$, $p < .05$. Younger drivers comprehended the signs from a greater distance (M=1370 feet, SE = 98.8) than did middle-aged or older drivers who did not differ from each other (M = 1053 feet, SE = 58.7; M = 1057 feet, SE = 85.5, respectively).

Night time comprehension stopping distance

We had complete data from 5 younger, 14 middle-aged, and 7 older drivers. Again, ANOVA showed only a significant effect of age. Again, younger drivers comprehended the signs at a greater distance (M = 1251 feet, SE = 66) than did middle-aged and older drivers, who did not differ from each other (M = 911 feet, SE = 39.5; M = 836 feet, SE = 55.9).

Day time pattern matching stopping distance

Complete data were available from 6 younger, 17 middle-aged, and 7 older drivers. The only significant effect in the ANOVA was for age, $F(2,27) = 6.79$, $p < .01$, though there was a trend ($p < .07$) for sign type. This time all age groups differed from each other with younger drivers successfully matching the sign at a greater distance (M = 1442 feet , SE = 97.2) than middle-aged drivers (M = 1073 feet, SE = 57.8), who in turn matched at a greater distance than older drivers (M = 993 feet, SE = 90.0). The trend was for fluorescent-sheeted signs to be matched at a greater distance (M = 1194 feet, SE = 46.4), than standard-sheeted signs (M = 1144 feet, SE = 53.2).

Night time pattern matching stopping distance

Complete data were available from 6 younger, 14 middle-aged, and 7 older drivers. ANOVA showed that the only significant effect was age, $F(2, 24) = 15.9$, $p < .01$. Younger drivers matched at a greater distance (M = 1242 feet, SE = 59) than middle-aged and older drivers, who did not differ from each other (M = 886 feet, SE = 38.7; M = 831 feet, SE = 54.7).

We also examined some additional radar-based measures: top speed, average speed, deceleration distance, but found essentially the same results as for stopping distance, the conventional measure for sign legibility.

Night Headlamp Beam Analysis

As mentioned in the methods section, drivers could select either low or high beams for their headlamps during the field study to approximate their normal night driving situation. We analyzed whether the beam setting influenced stopping distance. Given that drivers were not randomly assigned to low versus high beam conditions, this analysis is suggestive only and cannot lead to strong conclusions. We had too few cases with complete data to analyze by driver age group so collapsed across that variable. An ANOVA examining beam type (low, high), task type (pattern matching, comprehension), and sign sheeting (fluorescent, standard), revealed only a significant interaction between beam type and sign sheeting, $F(1,18) = 6.35$, $p < .05$. Figure 8 shows that fluorescent sheeting yields a slight advantage with low beam settings, whereas the opposite was true with high beam settings.

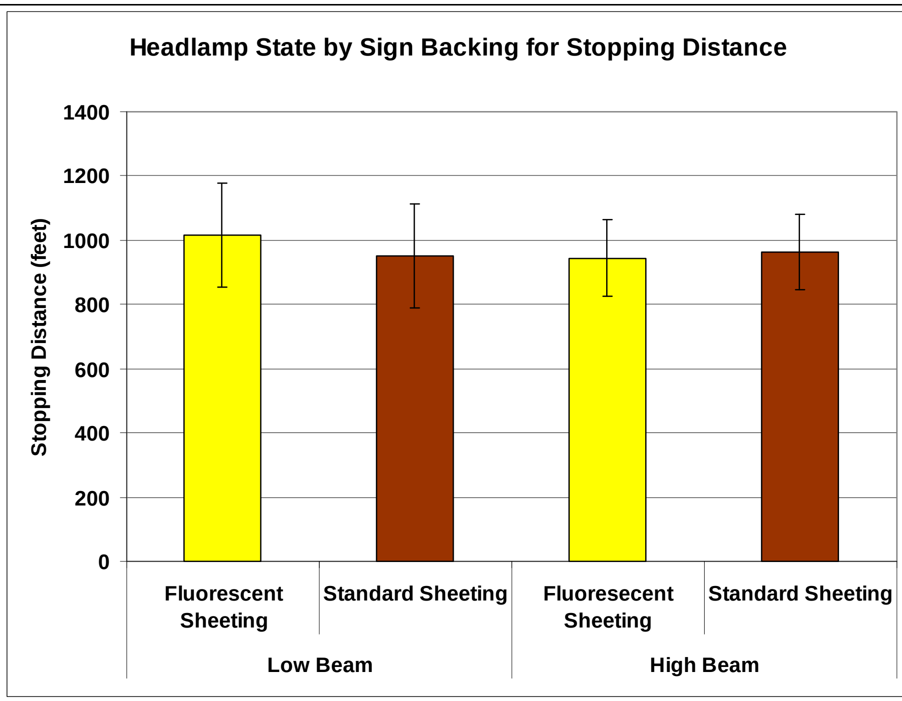


Figure 8. Influence of selected headlamp beam setting and sign sheeting on mean stopping distance, averaged across pattern matching and comprehension tasks. Error bars represent 95% confidence intervals for the means.

Despite the findings of occasional superiority for fluorescent sign sheeting over standard sheeting in the extant literature, we failed to find much of an advantage for that sign type, and when it occurred it occurred in daytime rather than night time (though in the latter case there was a trend for superiority for low beams versus high beams). Our study was underpowered because of significant attrition of participants. Nonetheless, age effects were quite robust even with the relatively small sample size. Older drivers, and occasionally middle-aged drivers, perceived signs, regardless of type, at lesser distances than younger drivers.

Correlations between lab and field performance

A main goal for this line of research is to assess whether inexpensive lab studies can predict successfully to more expensive field studies. Because we could not adequately simulate sign sheeting (fluorescent, standard) we average across those conditions in the field study. Similarly, because we could not choose a particular distance as most representative, we averaged across distance for the lab study. We first assess how the different averaged measures cohere. A factor analysis was run on accuracy and stopping distance for field measures and accuracy and response time for lab measures (both pattern matching and comprehension average measures). Three factors emerged (eigenvalues greater than 1), explaining 79% of the variance in the measures. The Varimax rotated factor solution is shown in Table 1. An oblique rotation method (oblimin) yields almost exactly the same solution.

Table 1. Rotated Component Matrix for Lab and Field Outcome Measures

	Component		
	1	2	3
Field_Day_Avg_Accuracy	.048	.054	.931
Field_Night_Avg_Accuracy	.201	.810	.046
Field_Day_Avg_Full_Stop	-.848	.283	.315
Field_Night_Avg_Full_Stop	-.779	-.065	.495
Lab.Comp.Avg.RT	.798	.092	.305
Lab.Comp.Avg.ACC	-.098	.847	-.184
Lab.PM.Avg.ACC	-.129	.690	.281
Lab.PM.Avg.RT	.956	.034	.149

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

The first factor represents strong positive loadings for the laboratory comprehension and pattern matching response time variables, and strong negative loadings for the field day and night stopping distance measures. Thus, there is some evidence that lab and field response time and stopping distance variables are measuring a common construct, namely, how quickly signs are apprehended. The second factor has strong positive loadings for field night-time accuracy and both lab comprehension and pattern matching accuracy scores and perhaps represents general visual acuity. The third factor shows a strong loading from field daytime accuracy, and moderate to strong loadings from field daytime and night time stopping distance, and lab comprehension response time, and may represent a cautiousness or response bias construct (willingness to respond quickly when uncertain).

Given that most of the research literature focuses on stopping distance, ignoring accuracy, we examine the scatterplots of lab response time (for comprehension and pattern matching tasks separately) and field stopping distance, separately for day and night stopping distance. Given missing data, we have n=21 to n=29 pairs of points in the graphs shown below. Thus, we have minimal power to find significance and the standard error for the correlation will be relatively high, though because we are positing a direction (negative) for the correlation, we can use one-tailed tests to increase sensitivity.

Figure 9 shows that there is a significant and strong prediction from comprehension response time in the lab to day time stopping distance in the field (averaging across comprehension and pattern matching tasks for fluorescent and standard sheeting), accounting for about a quarter of the variance in the field study measure. The relationship is negative because a longer stopping distance is associated with a shorter response time.

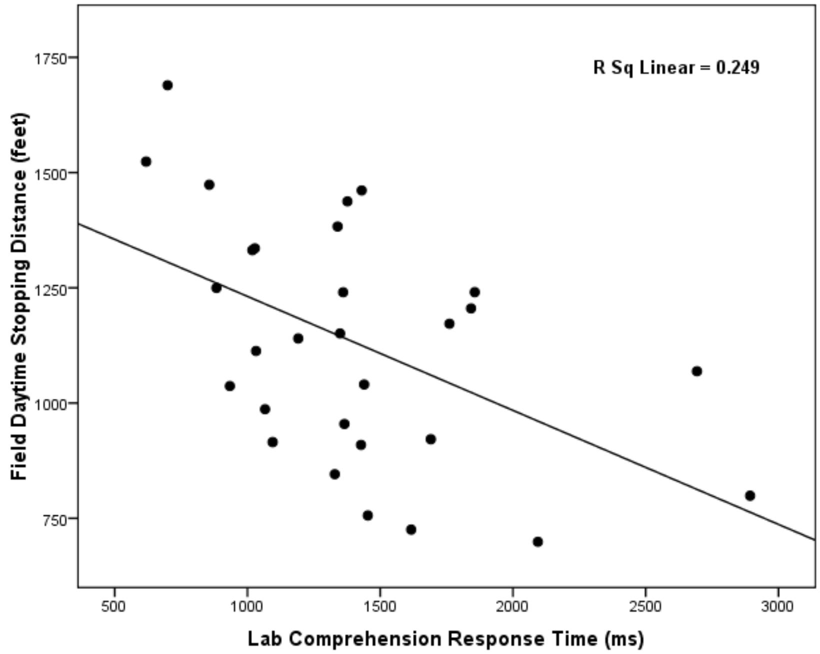


Figure 9 Scatterplot of field daytime stopping distance by lab comprehension response time. $r^2 = .25$.

An even stronger prediction of field performance is afforded by lab pattern matching as seen in Figure 10. Here nearly 40% of the variance in field performance is accounted for by the lab task.

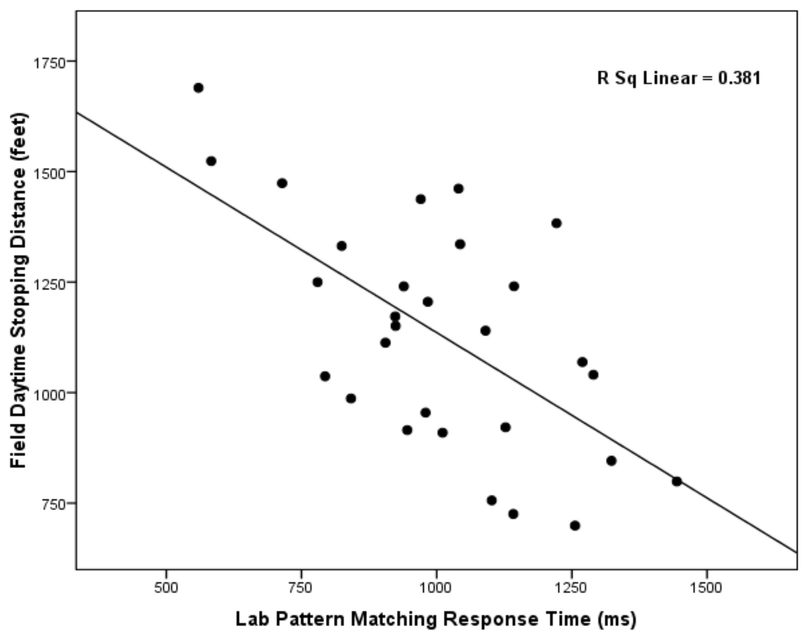


Figure 10. Scatterplot of field day time stopping distance by lab pattern matching response time. $r^2 = .38$.

Prediction from lab to field is equally impressive for night time performance as seen in Figures 11 and 12. Given our relatively small sample size, an outlier score (top panel, rightmost data point) is adversely impacting the relationship and the bottom panel shows the relationship without that data point.

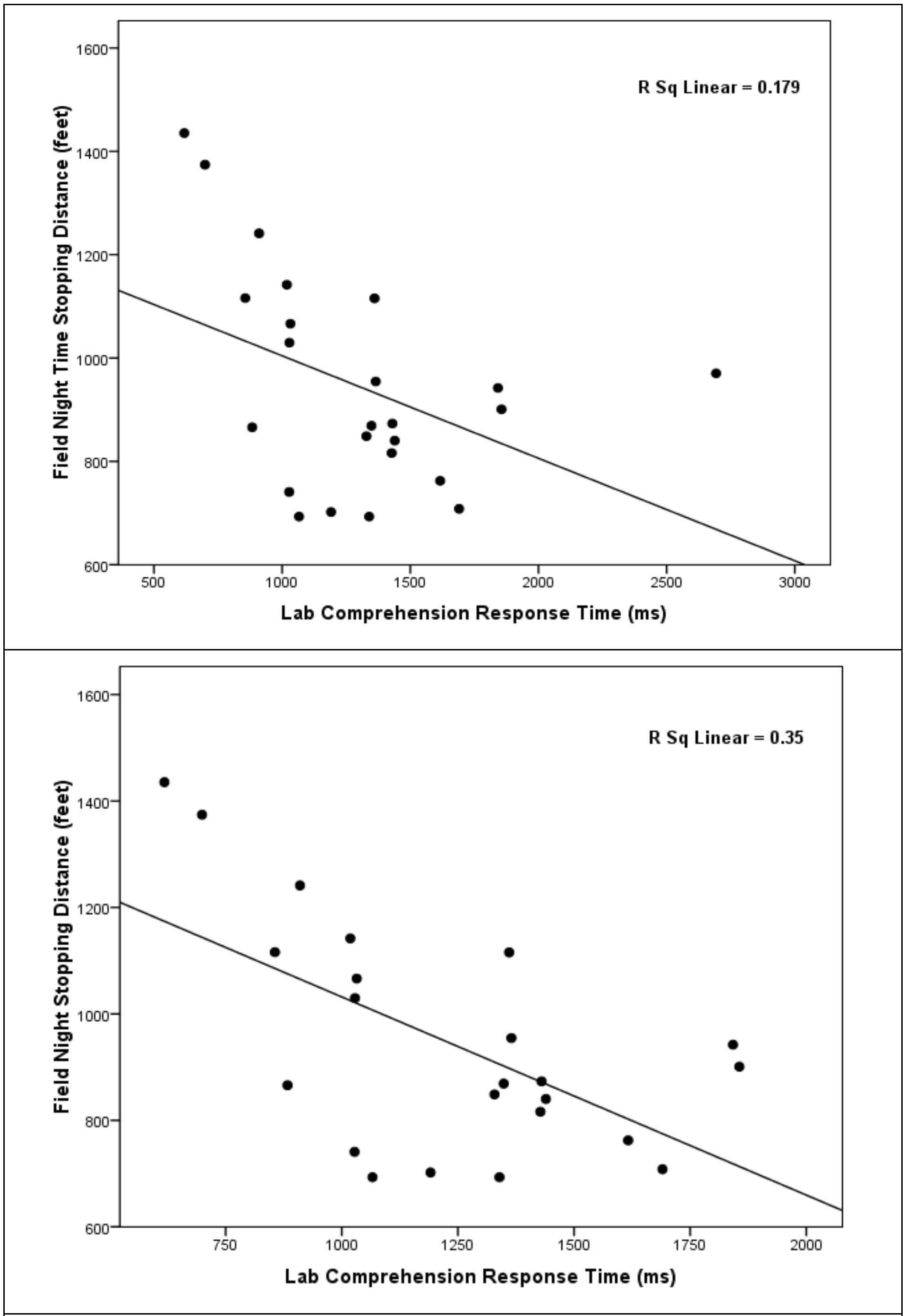


Figure 11. Scatterplots of field night time stopping distance by lab comprehension response time, with(r^2 = .18, top panel) and without (r^2 = .35, bottom panel) an outlier score.

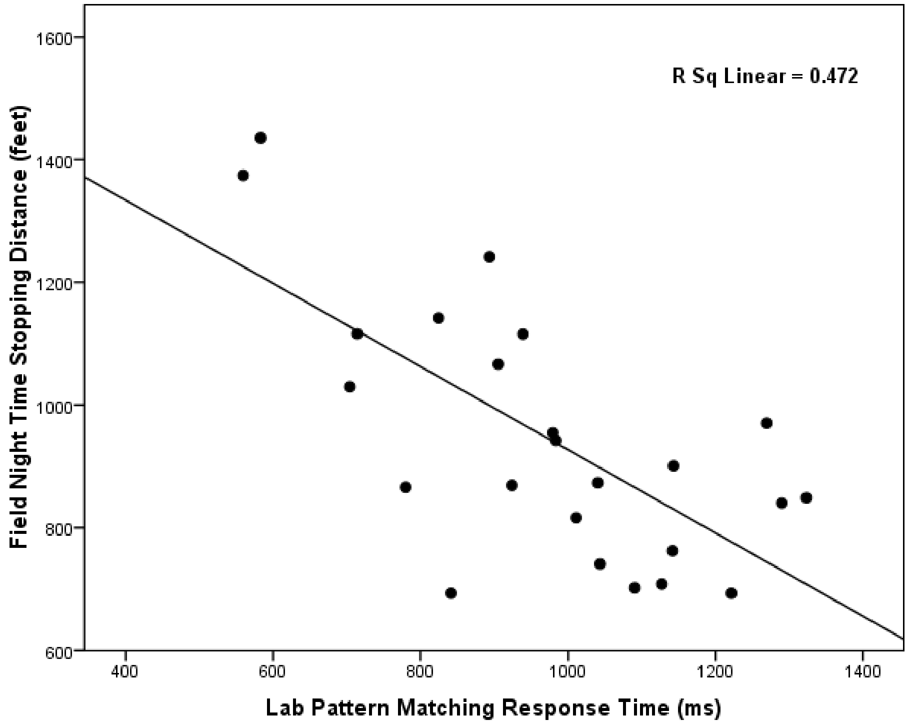


Figure 12. Scatterplot of field night time stopping distance by lab pattern matching response time. $r^2 = .47$.

This Figure indicates that we can account for nearly half the variance in field performance using the lab-based pattern matching response time measure. So, we have reasonable evidence that lab-based research studies, particularly those using pattern matching and comprehension response time, can accurately predict field stopping distance for warning signs.

Study 1 Conclusions.

In both studies there were robust effects of age group, with older drivers taking longer or moving closer to the warning signs before comprehending or matching them. Given that increased driver age magnified the effects of other variables (such as distance in the lab task), we recommend that older drivers be consistently incorporated into studies of traffic sign perception. They are the most sensitive to varying sign attributes.

Consistent with earlier findings in the literature we showed a small, though non-significant advantage for fluorescent sheeting over standard sheeting in the field study. There was a hint that the effect might be partially mediated by intensity of headlamps, with fluorescent sheeting doing better under low beams and perhaps worse under high beams. Thus, we would not recommend adopting fluorescent sheeted warning signs if their expense is significantly greater than that for standard sheeting. More research is needed to establish the relative efficacy of those types of signs under both low and high beam settings.

Our second goal for this set of studies was to determine if lab-based analogs of field-based tasks accurately predict field-based performance. We found that this is indeed the case for lab-based pattern matching (or comprehension) response time measures when predicting field-based stopping distance for warning signs. The lab-based pattern matching task seems to be the most sensitive one so we would advise the adoption of that procedure. The most popular field procedure is naming, rather than pattern matching or comprehension, so additional research may be needed to assess if the successful prediction holds for that field based task. However, the factor analysis of lab and field based measures give us some confidence that these procedures are tapping a common construct, warning sign perception.

Chapter 3.

Study 2: Visibility Factors for Traffic Signs

A street name sign can vary along many dimensions, such as font type (e.g., serifed, non-serifed), number of letters, size of letters, case of letters (uppercase, lowercase, mixed case), foreground and background color (e.g., white on green, black on white), and contrast ratio of foreground to background which will vary depending on ambient luminance (e.g., day versus night) and type of sheeting. Although there is considerable research on such factors for highway guide signs, there is less research with street name signs.

Garvey, Pietrucha, & Meeker (1997) showed an occasional advantage in 65+ year old drivers for 12.7 inch Clearview fonts over standard fonts, and inconsistent results for mixed case compared to uppercase letters using field studies during daytime and night time. Their first experiment using a recognition procedure presented a white on green sign with 3 words of length 5 or 6 letters, and the drivers, seated in the passenger seat of a test vehicle driven at ~10 mph, were to identify the target word in the set (indicating: top, middle, bottom for its position). In the second study, termed a legibility procedure, a single word was used, and drivers, seated as passengers, were asked to read the word. In the first study, mixed case was superior to uppercase font in some conditions. In the second study, all uppercase letters demonstrated a significant advantage for the distance at which the name was readable with Clearview D compared to Clearview condensed. One difficulty with interpreting the findings is that a relatively small sample of older drivers participated (n=12) so there was little power in the design to detect differences.

Forbes, Moscowitz, & Morgan,1950, cited in Garvey et al. (1997), apparently showed that mixed case was superior to uppercase font for highway guide signs in a recognition task, but not for a naming task. Smiley et al. (2001) used a paradigm of having drivers signal when they were to turn for a target street name in an on-road test of letter height legibility in downtown Toronto and found the larger letters, 20 cm (8 inch), clearly superior to smaller lettering. Uppercase letters are larger than lower case letters so we would expect a consistent legibility advantage for uppercase.

Guerrier & Fu (2002) studied middle-aged and older drivers' ability to name existing and specially prepared street name signs (ground-mounted name signs and advance street names) which varied in font type (Clearview, and Highway Series C, D, & E) and letter size (e.g., 4" to 12" heights) in a series of studies using roadways in the Miami area during day and night conditions. Drivers drove the routes in a test vehicle. In the first daytime study, using existing advance name and ground-mounted name signs, greater age was associated with decreased legibility distances, except for advance street name signs (the largest signs). In the second night time study using specially created signs, drivers followed a route around airport roads and identified 3 advance name (8 inch lettering) and 3 ground-mounted street signs (6 inch lettering), mostly between 6 and 7 letters long printed white on green backgrounds. The test car used high beams for this evaluation, and drivers were asked to name the street as quickly as possible. Clearview was superior to other fonts in terms of legibility distance for advance name signs but was judged equivalent to other fonts for smaller street name signs. In both studies, however, driver characteristics other than age -- performance on Trails-B, an attention/executive function test, and visual acuity -- were the primary predictors of legibility distance with age becoming non-significant in a multiple regression analysis. Although the sample sizes were greater here than in prior studies (n=51 in daytime, n=37 at night), the number of signs tested for the night time viewing condition was small (6) and the roadway conditions varied significantly for these signs (some well lit, some not well lit) and placements also varied considerably. It is important to try to hold constant the position of the sign in order to make good comparisons across other variables in case those variables interact with placement characteristics.

In a more recent study (Holick & Carlson, 2003), used a closed road system (airport taxiways and runways) to assess night-time legibility of Clearview and Series E fonts, with some variation in the case (upper, mixed) and size (16 inch highway, mixed case; 4 inch, 6 inch uppercase; 8 inch mixed case) for highway guide signs and road name signs. They tested three age groups (mean age of 27, 39, and 66 yr) with 10 participants/ age group in a test vehicle driven at 35 mph, with participants required to name the sign in order to assess legibility distance. They tested at night with car headlamps set to high and low values for a low beam setting by a customized pulse-width modulation dimmer (13, 6 cd/m^2 luminance assessed from a blank sign at a distance of 640 feet). They used a broad range of signs (138) but to avoid having drivers see the same signs randomized the ordering so that any driver saw between 2 and 4 repetitions of the 3 variables (alphabet, background sheeting, legend sheeting) across the 3 sign types (shoulder-mounted guide, destination/distance, common road sign). For freeway signs (6 letters), legibility distances were greater with Clearview font compared to series E; they decreased with age; legibility distance increased with luminance; there was also a luminance by sheeting interaction showing that as retroflectivity contrast ratio increased, legibility distance improved more for low luminance than high luminance conditions. Microprismatic sheeting coupled with Type III background performed best. Similar results were obtained for the destination and distance signs (5-10 letters) as well as the country road signs (6-10 letters), though overall statistical tests (ANOVA) were not described. Unfortunately, although this study use some of the most sophisticated methodology of those reviewed, it did not vary case (upper, mixed) with other variables, or vary number of letters in sign names systematically (or report results based on letter number), and the small sample size for age groups and the use of a relatively young older group meant that there was little opportunity to detect interactions of age with other variables, particularly for older drivers.

A weakness for the Holick study, shared by other studies reviewed, is that accuracy data were not reported for the tasks. It may be the case that all signs were named correctly in which case there is no difficulty associated with interpreting the obtained legibility distances. However, there is a well known problem, known as speed-accuracy trade-off, for response time and accuracy. Depending on the criterion used by the driver, they can obtain longer legibility distances at the cost of making more naming or matching errors, or they can approach perfect accuracy by delaying their decision until they are very close to the sign. Older adults have been shown to be more conservative (preferring accuracy to speed) in laboratory research studies (Fisk et al., 2009). Thus, it is wise to record accuracy as well as legibility distance. In laboratory studies this is usually done by examining both factors (speed, accuracy).

In our field and laboratory studies, we varied factorially the following sign characteristics: length of name (6, 8, 10 letters for the lab portion; 8, 10 for the field portion), case (upper versus mixed), and day versus night viewing for the field study. We chose length 8 and 10 character words for the field study because they would represent the most difficult legibility conditions for older drivers at night. (Time to process a word is usually an increasing function of word length.) As shown in Figure 11, mean name length in Leon County is 8 letters. Prior research investigating this factor used shorter name lengths so we do not have much empirical evidence for the effect of longer street names. We assessed matching distance for younger, middle-aged, and older drivers in the field study, and response time for the lab study. Accuracy in the task was also assessed. For the field task, drivers completed sign identification tasks in their own vehicles.

In the laboratory variant of the sign matching study we assessed both speed and accuracy for sign perception and varied word length, simulated distance to sign (near, far, using smaller and larger signs) to enable us to look at trends for shorter than average sign word lengths. A subset of participants in the lab study also participated in the field study, given refusals to return for the field study and some data losses during field data collection. Thus, we have a larger sample size for the laboratory study: 52 younger, 36 middle-aged, and 24 older drivers, with more power to detect small effect sizes.

Method

Study 2 Laboratory Task.

Design

Study 2 Lab examined correlations across laboratory and field with a simulation of the street name matching task used in Study 2 Field. A 3(age: younger, middle, older) x 2(distance: ~ 35 ft., ~70 ft) x 2(letter-case: mixed, uppercase) x 3(word-length: 6, 8, 10) mixed factorial design was used with age as a between-subjects variable, and distance, letter-case, and word-length as within-subject variables. Dependent variables were reaction time and accuracy.

Participants

Fifty-two younger (M = 25 yr), 36 middle-aged (M = 58 yr), and 25 older (M = 73 yr) drivers participated.

Materials and Procedure

Pattern matching Street Names

The task was very similar to the pattern-matching and comprehension tasks from Study 1 Lab, and was administered using the same equipment. Forty-eight novel street names were coined that were either six, eight, or ten letters in length. A counterpart to each name was created by changing one vowel in every name, resulting in a total of 96 names. Orthographic conventions were followed when vowels were changed to minimize spellings that would appear odd or unusual.

Trials began with presentation of a street name in the center of the monitor. The street name was printed in Courier 18-point font in all lowercase letters, and occupied roughly 0.5 degrees of visual angle per letter. When the participant had correctly pronounced the street name, the experimenter pressed a button that advanced an image of a street name sign in the center of the screen. The sign appeared in one of two sizes, corresponding to distances of ~35, and ~70 ft. Participants decided if the street name on the sign matched the initial street name, and indicated their answer by pressing one of two keys labeled “same” and “different.” Participants were instructed to make this forced choice as quickly as possible while sustaining accuracy.

The task was administered in two counterbalanced blocks—one with near sign and one with far signs. Within either block, all 96 trials occurred in a fixed pseudo-random order with each initial street name displayed twice, and followed by its match and mismatch one time each. Prior to the task, participants practiced on four trials and received feedback of their accuracy. Feedback was not provided thereafter.

Results.

Accuracy

Accuracy on the sign matching task was quite high on average (94%). Age group was not a significant predictor of accuracy, nor did age interact with other conditions. Thus, again, we are in a good position to assess speed of performance because age groups did not vary on accuracy. Given the large number of variables in the analysis, we used a multivariate test followed by examining univariate (within-subjects) tests for this mixed between/within design. The statistically significant main effects were for case and number of letters, but these were moderated by significant interactions of distance x number of letters ($F(2, 220) = 3.35$, $p < .05$), and case by number of letters ($F(2, 220) = 81.8$, $p < 01$. The distance (sign size) by number of letters interaction, shown in Figure 13 indicates that distance to sign (simulated by sign size) had a differential effect for long sign names. The 10-letter names were slightly, though significantly more accurately identified when the sign was closer. For both 8 and 6 letter signs, accuracy was uniformly high for near and far distances, though accuracy is somewhat greater for 6-letter names.

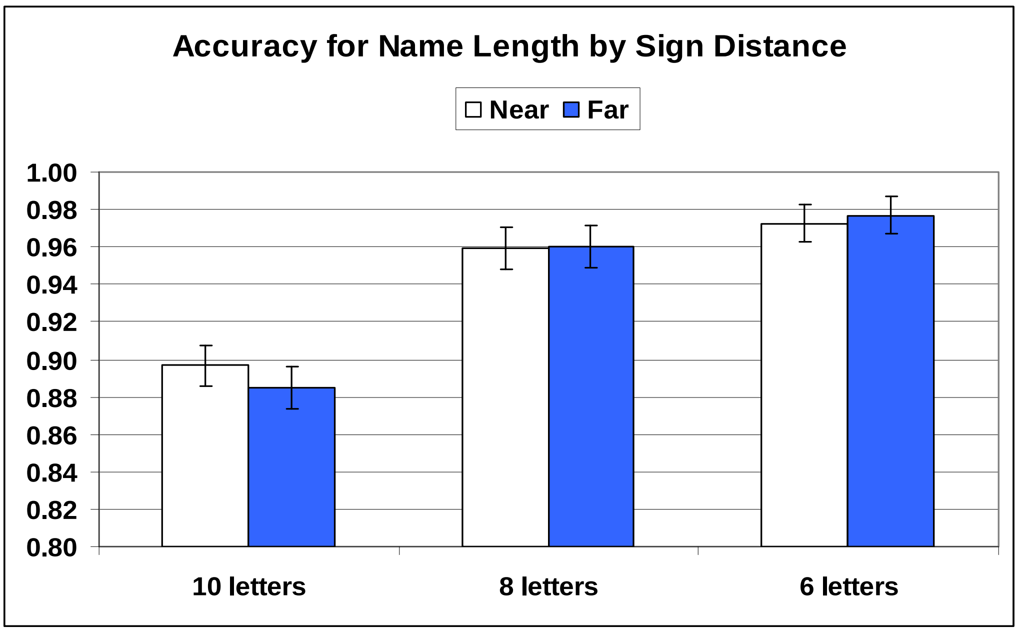


Figure 13. Mean proportion correct by sign distance and number of letters in the sign. Standard error bars represent the 95% confidence interval for the mean.

The case by name length interaction is show in Figure 14. There is an accuracy advantage for uppercase lettering for the longest street name (10 letters), otherwise there is equivalence for uppercase versus mixed case lettering.

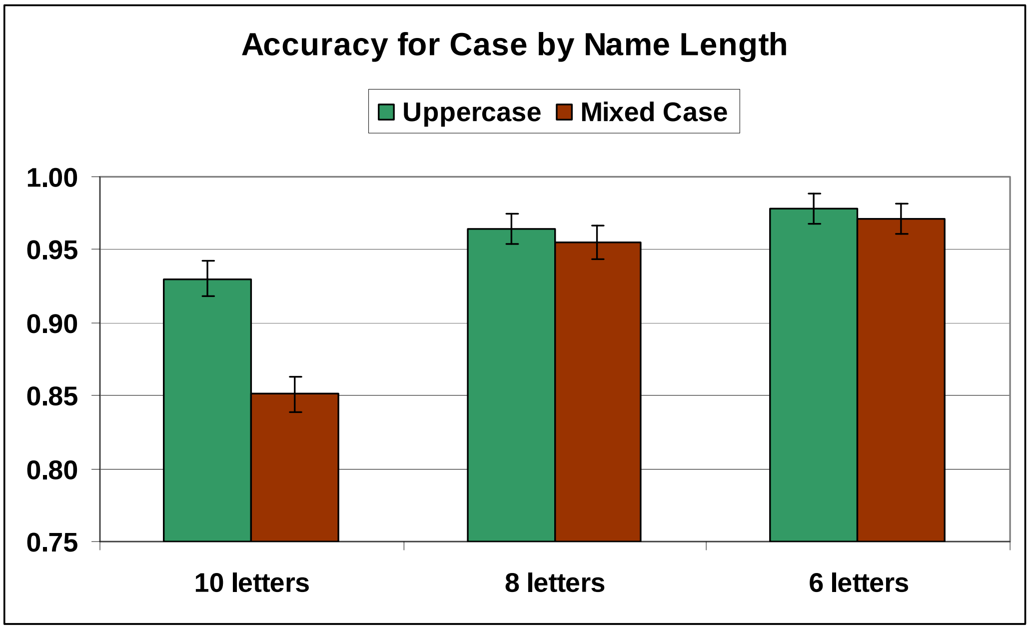


Figure 14. Mean proportion correct by letter case and number of letters in the sign. Standard error bars represent the 95% confidence interval for the mean.

In summary, accuracy is very good for all but the most difficult judgments: the 10-letter street names and signs at simulated far distances or those of 10 letters in mixed case lettering. Laboratory accuracy is also somewhat higher than field accuracy (discussed below), suggesting that laboratory studies will be slightly positively biased when accuracy of perception of street signs is concerned.

Response time

For response time, we have numerous significant effects, but these were moderated by the highest order interaction: a 4-way interaction of distance by number of letters by case by age group ($F(4, 220) = 2.43$, $p < .05$). We first present in Table 2 mean response times, then we present Figure 15 to try to explicate this interaction.

Table 2. Four-way Interaction of Age, Distance, Case, and Word Length

Age Group	Distance	Case	Word Length	Mean	Std. Error
Younger	near	upper	10 letters	1246	63
			8 letters	1113	51
			6 letters	1012	44
		mixed	10 letters	1290	59
			8 letters	1120	50
			6 letters	988	41
	far	upper	10 letters	1212	69
			8 letters	1126	54
			6 letters	1032	51
		mixed	10 letters	1256	75
			8 letters	1113	61
			6 letters	995	47
Middle	near	upper	10 letters	1447	76
			8 letters	1280	62
			6 letters	1115	53
		mixed	10 letters	1427	71
			8 letters	1248	60
			6 letters	1080	49
	far	upper	10 letters	1514	83
			8 letters	1331	65
			6 letters	1187	61
		mixed	10 letters	1540	90
			8 letters	1372	73
			6 letters	1183	56
Older	near	upper	10 letters	1721	91
			8 letters	1434	74
			6 letters	1209	64
		mixed	10 letters	1611	86
			8 letters	1444	72
			6 letters	1196	59
	far	upper	10 letters	1712	100
			8 letters	1435	78
			6 letters	1351	73
		mixed	10 letters	1798	109
			8 letters	1569	88
			6 letters	1275	67

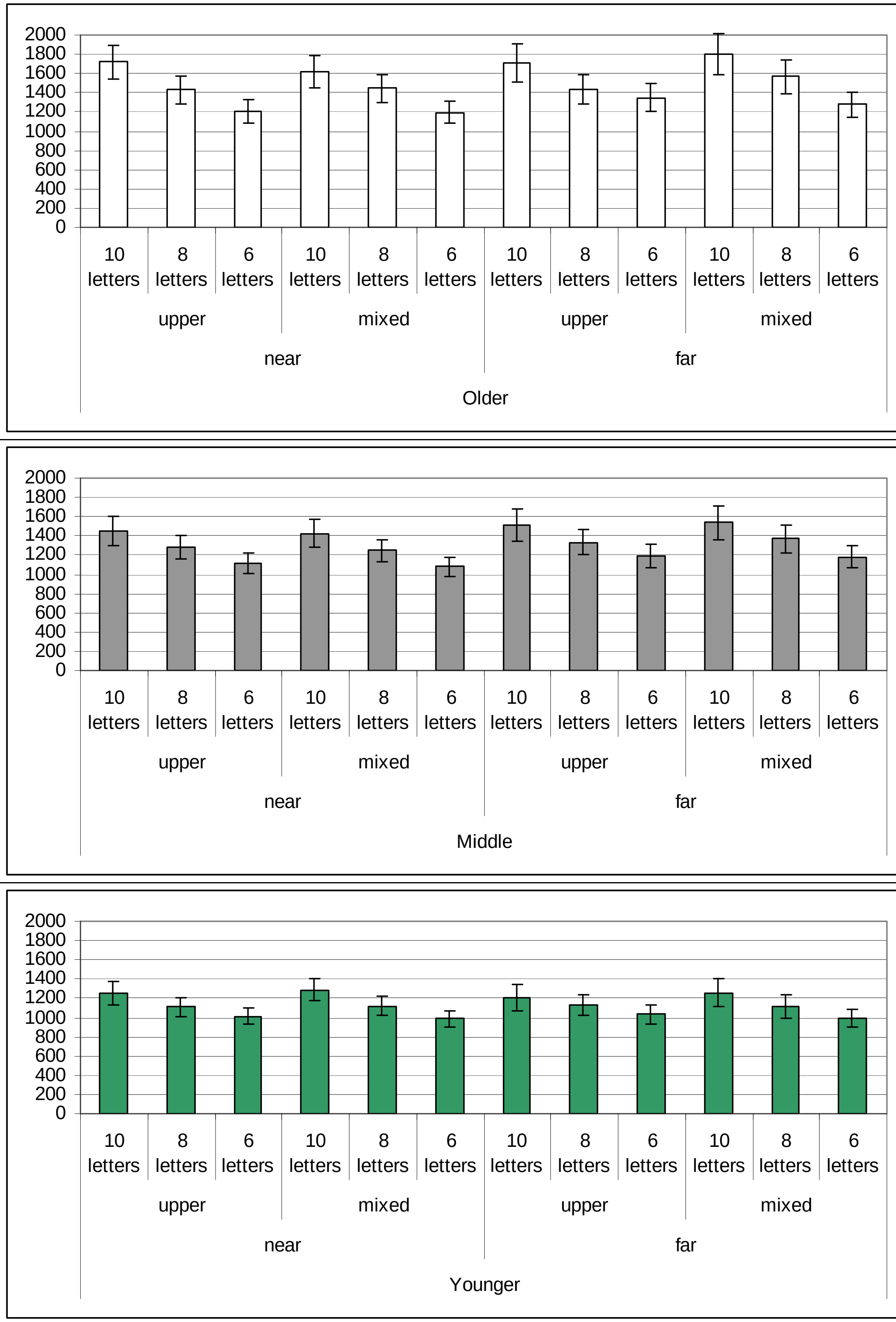


Figure 15. Mean response time in milliseconds as a function of age group (older, middle-aged, and younger drivers), number of letters, sign distance (size, simulating near, far)), and case (uppercase, mixed case) Standard error bars represent the 95% confidence interval.

A major factor is age, with older drivers much slower than middle-aged drivers who in turn are slower than younger drivers to make a match or mismatch decision about a sign. Younger drivers respond more quickly to 6 than 8 than 10 letter signs, but are relatively unaffected by sign case or sign distance. Middle-aged drivers are affected similarly by number of letters in the sign, but they also show sensitivity to sign distance (faster to near than far signs), though not much sensitivity to sign case. Older drivers show increased sensitivity to all 3 factors. They have a much lengthened response to longer street name signs. They show some moderation of this effect by sign distance, with somewhat greater increase in response time to distant compared to near signs. As well, they are the only age group showing sensitivity to letter case. For distant signs they tend to respond more quickly to upper case signs, though not for near signs. The easiest way to think about this interaction is that it shows that older drivers are the most sensitive to visual factors such as length of sign, case, and distance. Thus, minimizing the number of letters to be processed and increasing the size of letters (case, distance effects) helps them the most.

Study 2 Field Task.

Design

Study 2 Field examined the legibility of street name signs in a 3(age: younger/middle aged/ older) x 2(letter-case: mixed/capitals) x 3(word length: day/night) mixed design. Age was a between-subject factor and letter case and word length were within-subject factors.

Participants

Twenty-seven younger (M = 26 yr), 18 middle-aged (M = 59 yr), and 17 older (M = 73 yr) drivers were recruited to participate in the study.

Materials and Procedure

Although Study 2 Field was conducted in a different location than Study 1 Field, many of the same materials were used. What follows are the additional materials used in Study 2.

Signs

The signs used for this study are as follows: Camerlingo, Hackmatack, Orangeburg, Sayreville, Steelman, Nocturne, Earlsdon, Overhale, Mandelhall, Zancadilla, Heidelberg, Pettychaps, Thurston, Paddocks, Lollipop, and Isospore. For further information and specifics on each sign, see Appendix B.

Sign Changing Apparatus

A new sign changing apparatus was designed for Study 2 field that was capable of displaying signs at the height of 17 feet (see Figure 16). This device, made of all aluminum, was mounted four feet into the ground for stability. The main legs of the device formed an axle towards the top of the legs. The internal component was nearly 19 feet tall and rotated around the axle to lower the top end where the sign was mounted down to the ground for easy changing. The signs were held in place by brackets and stops. The device was then raised with the aid of ropes, pulleys, and counter-weights.

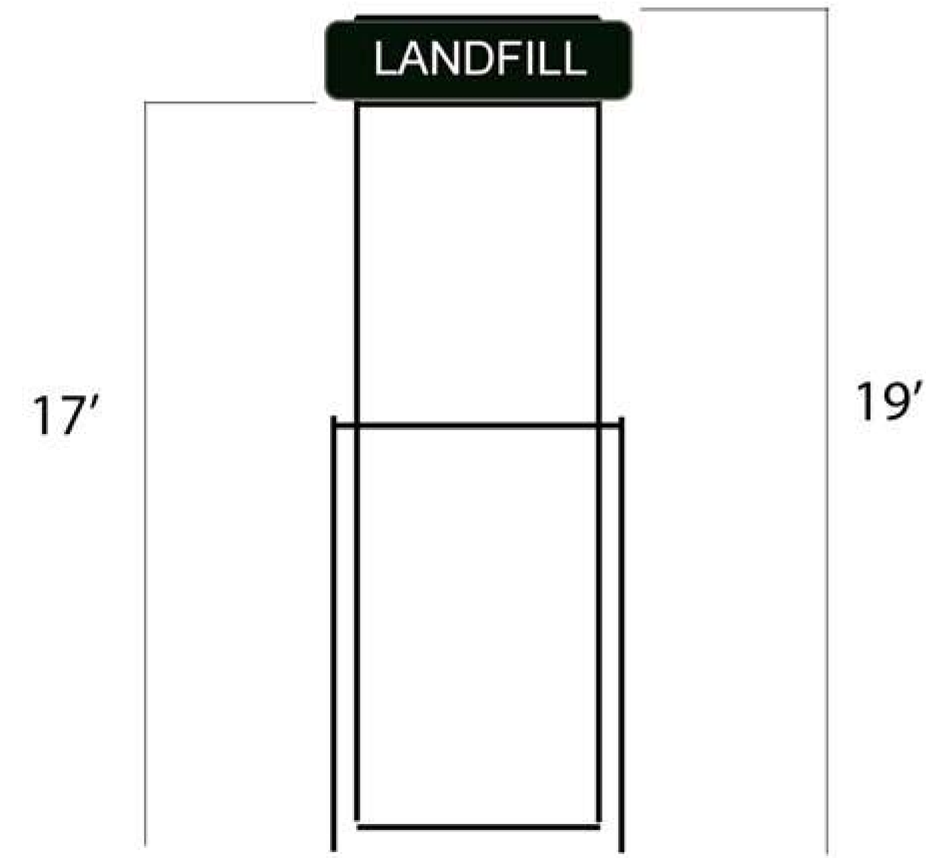


Figure 16. Diagram of Task 2 sign changing apparatus.

Procedure. The procedure was very similar to the comprehension and pattern matching tasks used in Field Study 1. At every trial participants were shown a street name in Clearview font on a laptop computer and instructed to drive forward at a comfortable speed until they could determine with 100% confidence whether or not the name on the street sign at the end of the track matched this street name. The street names matched for half of the 16 trials and did not match for the other half. Once again, distance and speed were measured with a radar gun and data were recorded on a laptop computer.

Results.

Given data loss (radar gun losses) and drop-out, we have complete data on a maximum of 26 younger, 17 middle-aged, and 17 older drivers for the day time task and 25, 18, and 16 respectively who continued on to the night task.

Accuracy

In this analysis we looked at whether the decision was correct for the matching judgment (sign matched, sign did not match). We found that the only statistically significant effect was for word length, $F(1, 54) = 5.66$, $p < .05$, with greater accuracy on 8-letter signs compared to 10-letter signs (91% vs. 87%) irrespective of age, day versus night, and upper versus mixed case for lettering. Thus, all age groups worked to the same accuracy criterion enabling us to draw stronger conclusions from the stopping distance data. Two trends are worth mentioning. First, there was a marginal (p=.057) interaction effect of session (day, night) by case (upper, mixed) by age group. For the middle-aged and older drivers accuracy was lowest mainly in the night condition with mixed case signs (dropping to 83% for middle-aged drivers, and to 84% during the day and 87% at night for the older drivers). Such results are suggestive that uppercase only signs would have a slight advantage, particularly at night, for older drivers, compared to mixed case ones. Second, there was a marginal ($p = .068$) effect of session. Not surprisingly, accuracy was marginally better during the day than at night (90% vs. 87%). Somewhat surprisingly, although average accuracy at night was significantly correlated with average accuracy during the day, the relationship was not that strong: $r = .27$. A multiple regression to assess whether age or gender might have been mediators of this weak relationship showed that this was not the case: $F (3, 53) = 2.14$, p >. 10.

Stopping distance to sign

In this analysis we examined where the driver was relative to the sign when they announced their decision about the sign (match, mismatch). Longer distances indicate better performance. The significant effects were attributable to age group ($F (2, 53) = 7.14$, $p < .01$), session ($F (1, 53) = 64.2$, $p < .01$), case ($F (1, 53) = 102$, $p < .01$, and word length ($F (1, 53) = 4.39$, p < 05). The session by case interaction was also significant: $F (1, 53) = 6.26$, $p < .02$). The age group effect can be seen in Figure 17, showing that middle-aged (365 feet) and older drivers (341 feet) had to drive much closer to the sign before making their match/mismatch decision than did younger drivers (443 feet).

Figure 17. Driver Age Group effect for stopping distance to sign. Error bars represent 95% confidence intervals for the means

Not surprisingly, people drove closer to the sign at night (338 feet) than during the day (428 feet). Similarly, they drove closer to the sign for 10 letter street names (380 feet) than 8 letter names (385 feet). The session by case interaction indicated that upper case signs were advantaged compared to mixed case signs more during the day than at night, as seen in Figure 18.

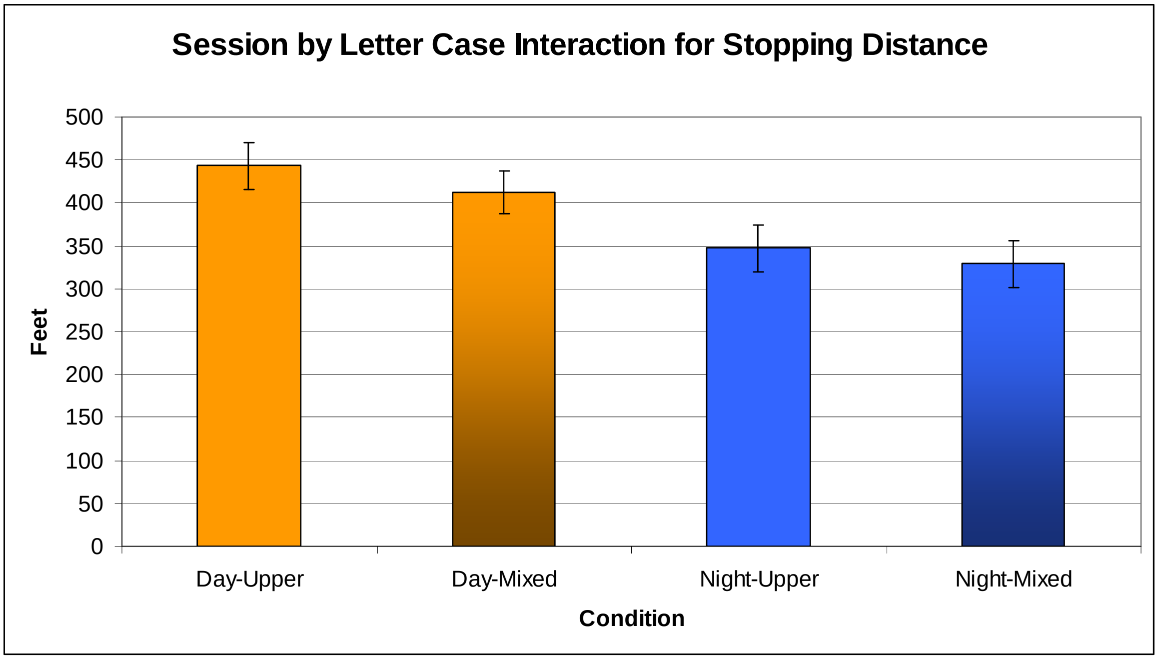


Figure 18. Session by Letter Case interaction for stopping distance. Standard error bars represent 95% confidence intervals.

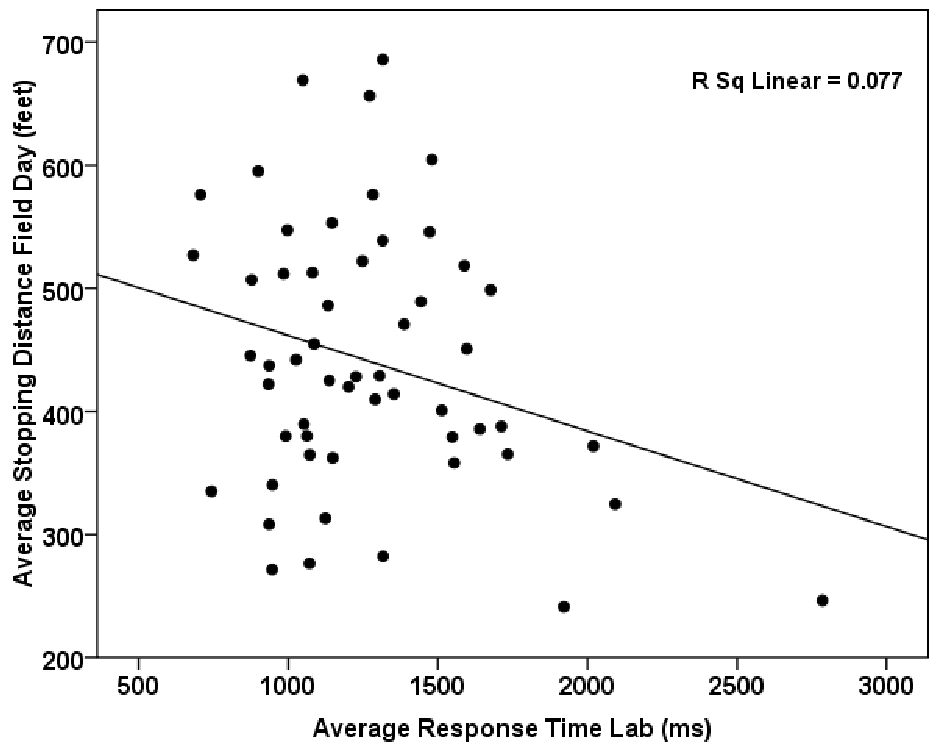
The greater advantage for upper case lettering on signs during the day may reflect a more conservative approach taken by all age groups at night for this task, where they chose to approach the sign very closely before making their decision irrespective of case of letters.

Based on the combined results for accuracy and stopping distance, we would argue that upper case street names will improve driver performance when trying to find unfamiliar street names. The results also indicate that use of shorter street names (such as abbreviations) may improve sign comprehension processes, though formal testing of comprehension of abbreviated names would be required.

Comparison between Lab and Field Performance

It is worthwhile to compare lab and field performance averaging across letter lengths and font types (upper, mixed case). In Figure 19 we show the relationship between average lab response time and average field stopping distance for day and night conditions. Note that we are conducting a conservative test of lab to field relationships by averaging across experimental conditions. For instance, lab uppercase 10 characters far distance response time is more strongly correlated with field stopping distance for uppercase 10 character day time signs (r^2=.12).

The only significant relationship is between lab performance and field performance during the day (not at night), though the two lab-field correlations are not significantly different from each other. There is a strong relationship (r = .70) between stopping distances during the day and at night.



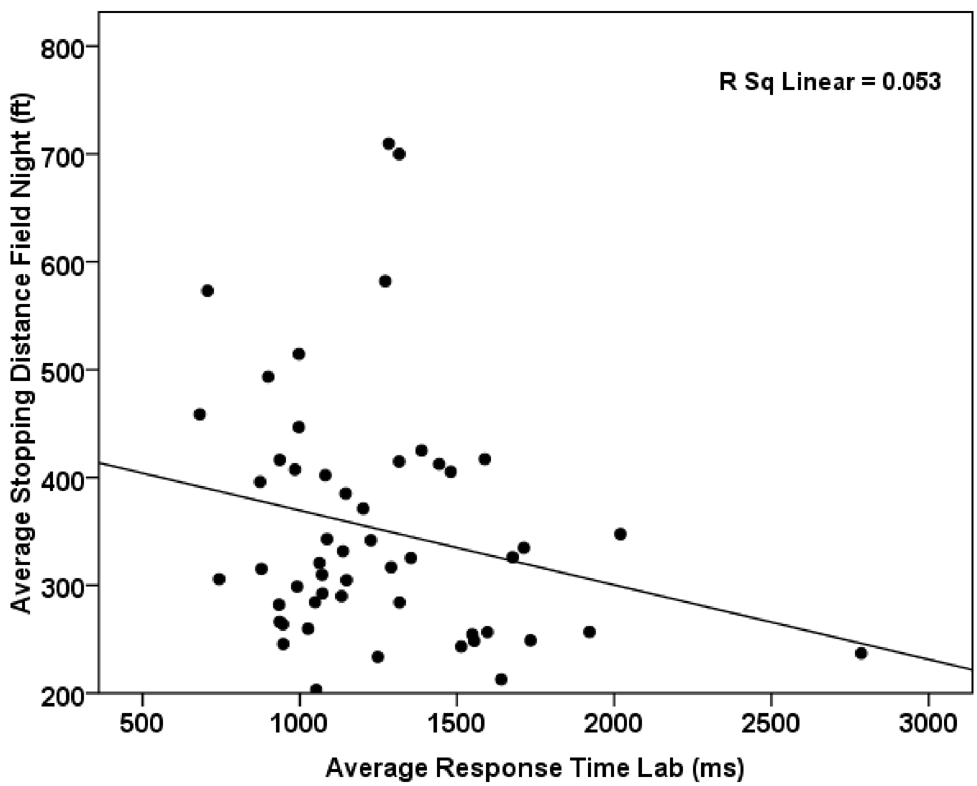


Figure. 19. Mean field stopping distance (ft) during the day averaged across case and 8, 10 letters (top panel) and at night (bottom panel) by mean lab response time (ms) averaged across 6,8,10 letters and mixed, uppercase fonts

So, average lab performance does not predict average field performance very well using these pattern matching procedures. In short, there is a need for field studies to supplement findings in lab studies when assessing legibility factors for street signs.

Study 2 Conclusions.

Uppercase street signs should provide the best legibility distances for the general driver population assuming that their task is to locate a particular street name (e.g., unfamiliar street name). The advantage is mainly found for older drivers, particularly for longer street names. There is consistent support for this finding in both field and laboratory experiments using both legibility distance (field) and response time (lab) measures as well as some indication in accuracy data (which tended to be uniformly high). The weak correlation between outcome measures across study types indicates that field studies are needed to validate the findings in lab studies.

Caveat

We did not evaluate the driving response to having street sign names sooner with uppercase lettering. Thus, our study is limited to the conspicuity aspects of sign perception and not to decision making processes that are downstream to perceptual ones.

Chapter 4.

Study 3: Symbols vs. Word Messages for Signs

In a series of laboratory studies (Dewar et al., 1997; Kline, Ghali & Kline, 1990), drivers were asked to identify both word and symbol (icon) traffic signs. There were fairly consistent results with legibility distances for older drivers much shorter than for younger drivers and symbol signs could be improved significantly in legibility when they emphasized low spatial frequencies rather than high spatial frequencies for differentiating their internal elements. However, there has been significant controversy over whether laboratory studies that carefully control visual conditions on a computer screen generalize effectively to road conditions where there is visual clutter in the driving environment and in which drivers are likely to be stressed, particularly older drivers (Zwahlen & Schnell, 1999).

In our studies we examined symbol versus word message sign comprehension distances and included both lab and field studies to address concerns about generalizability. The field study used both day and night viewing conditions. The lab study examined both comprehension and pattern matching tasks.

Method

Study 3 Laboratory Task.

Design

The laboratory study was designed to examine correlations across laboratory and field using similar tasks. A 3 age (younger, middle, older) x 2 sign type (word or symbol) mixed factorial design was used with age as a between-subjects variable, and sign type as within subject variables. Dependent variables were reaction time and accuracy.

Participants

Nineteen younger (M = 26 yr), 21 middle-aged (M = 57 yr), and 24 older adults (M = 72 yr) participated in the lab study.

Materials and Procedure

The materials and procedure for this study were almost identical to those used in Studies 1 and 2 Lab. The procedure is most similar to the Study 1 Lab comprehension task. The only difference is that the present task used images of the signs from Study 3 Field (size equivalent to a 20 foot distance) and they appeared in the center of the screen instead of the right-hand side of the screen. Each trial began with a 3 second presentation of one of the statements from Study 3 field in Courier 18-point font in the center of the screen. After each statement an image of a sign appeared in the center of the screen in one of the three sizes/viewing distances. Once again the goal was to indicate whether or not the statement and sign were semantically similar or different.

Signs

Signs used in this task are seen in Figure 20.

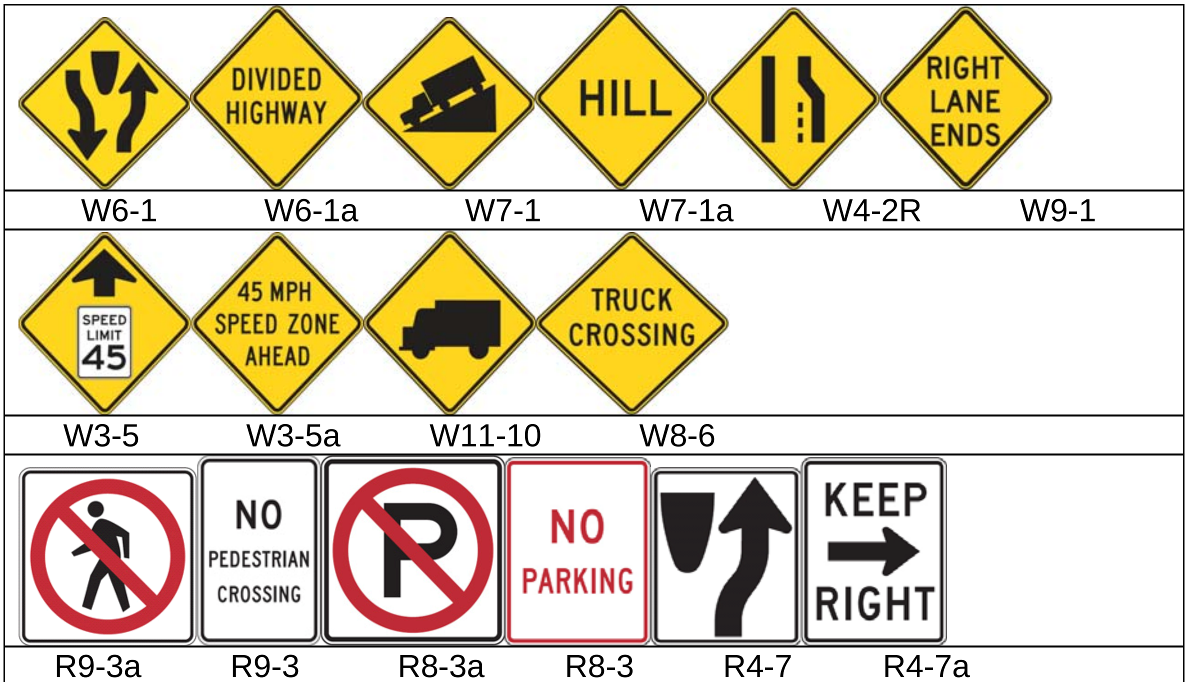


Figure 20. Signs used in the task 3 laboratory task.

Results

Study 3 Lab.

We had complete data from 19 younger, 19 middle-aged, and 24 older drivers.

Comprehension Measures

These measures tapped whether the drivers were able to comprehend the meaning of a sign from a description of it in words.

Comprehension Accuracy

The ANOVA showed significant main effects for distance (simulated by size of sign if viewed at 20 feet, 200 feet, 500 feet), $F(2, 118) = 94.1$, $p < .01$, sign type (symbol, word) $F(1, 118) = 5.71$, $p <. 05$, and age group, $F(1, 59) = 7.43$, $p < 01$. However the main effects were moderated by significant interactions for age group x distance, $F(4, 118) = 3.37$, $p < 05$, and distance x sign type, $F(2, 118) = 25.1$, $p < .01$. Those interactions can be seen in Figure 21 and 22.

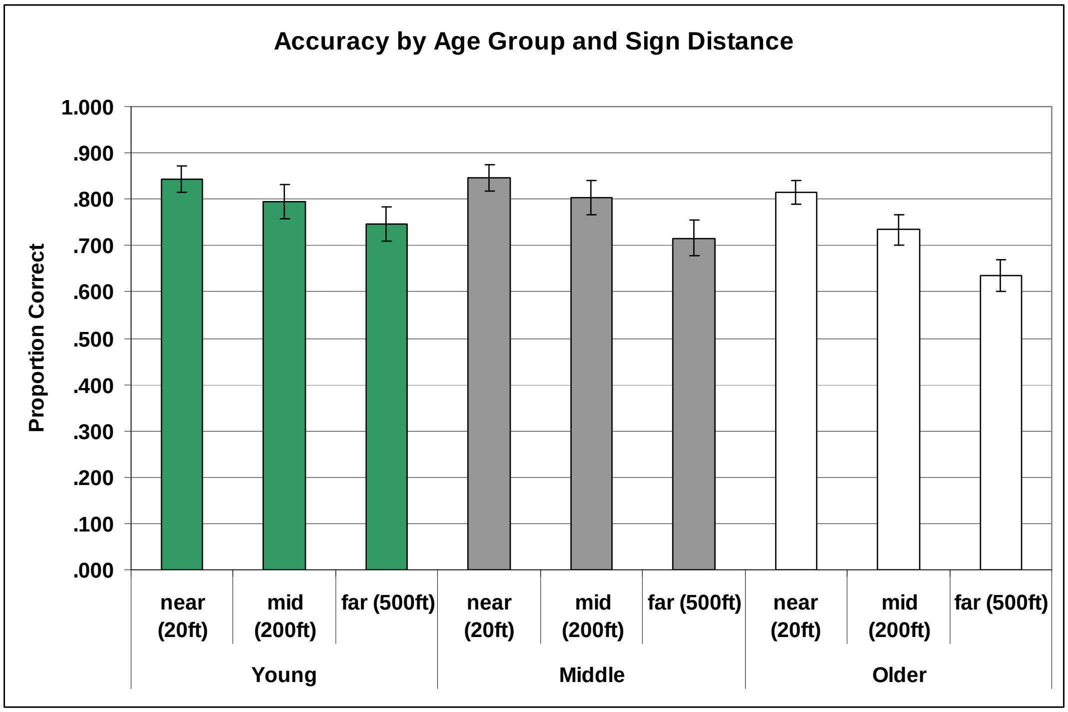


Figure 21. Mean proportion correct comprehensions by age group and sign distance. Error bars represent 95% confidence intervals for the mean.

The age x distance interaction indicates that the slope of the distance effect (less accuracy for greater sign distances) is steeper for older drivers than for the other two age groups. Comprehension accuracy is above 80 % for the near distance (20 feet) but drops off most sharply at the most extreme distance (500 feet).

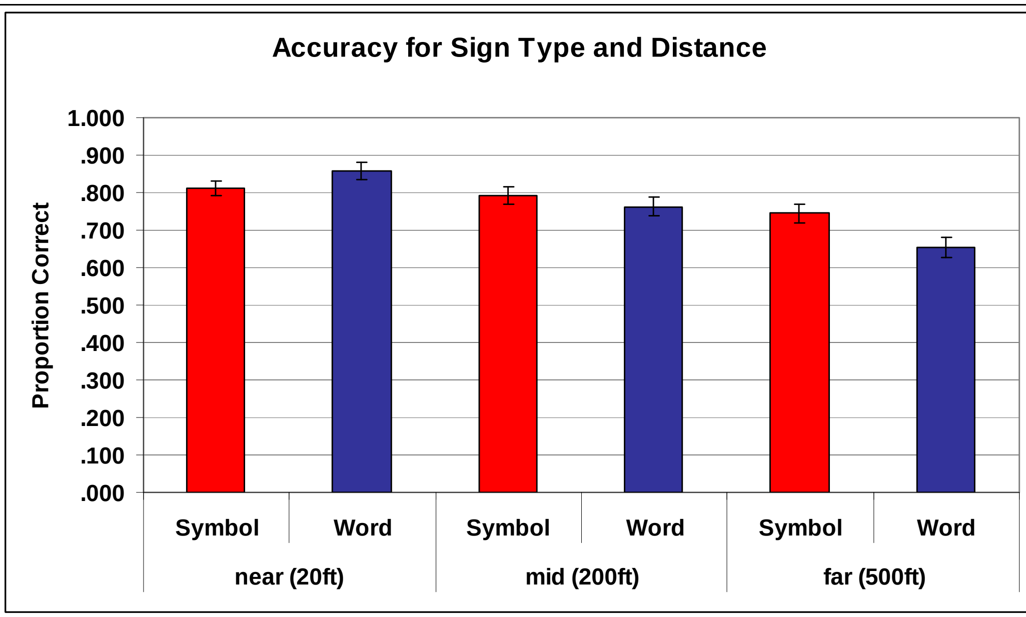


Figure 22. Mean comprehension proportion correct judgments by age group and sign type. Error bars represent 95% confidence intervals for the mean.

The sign type by distance interaction shows that symbol signs are recognized better than word signs at long distances (200, 500 feet), whereas the reverse is true for the shortest distance (20 feet).

Comprehension Response Time

Because accuracy varied interactively with age and distance as well as distance and sign type, caution is necessary to interpret response time. We assessed effects for correct trials only.

ANOVA showed significant effects of age group, F (1, 59) = 4.36, p < .02, distance, F (2, 118) = 19.2, p < .01, sign type, F (1, 118) = 46.6, p < .01, and a significant sign type x distance interaction, F (2, 118) = 6.71, p < 01. Younger drivers responded the quickest (M = 1667 ms, SE = 194), with middle-aged drivers responding more slowly (M = 2351 ms, SE = 194), and older drivers responding the slowest (M = 2359 ms, SE = 173). The interaction is shown in Figure 23.

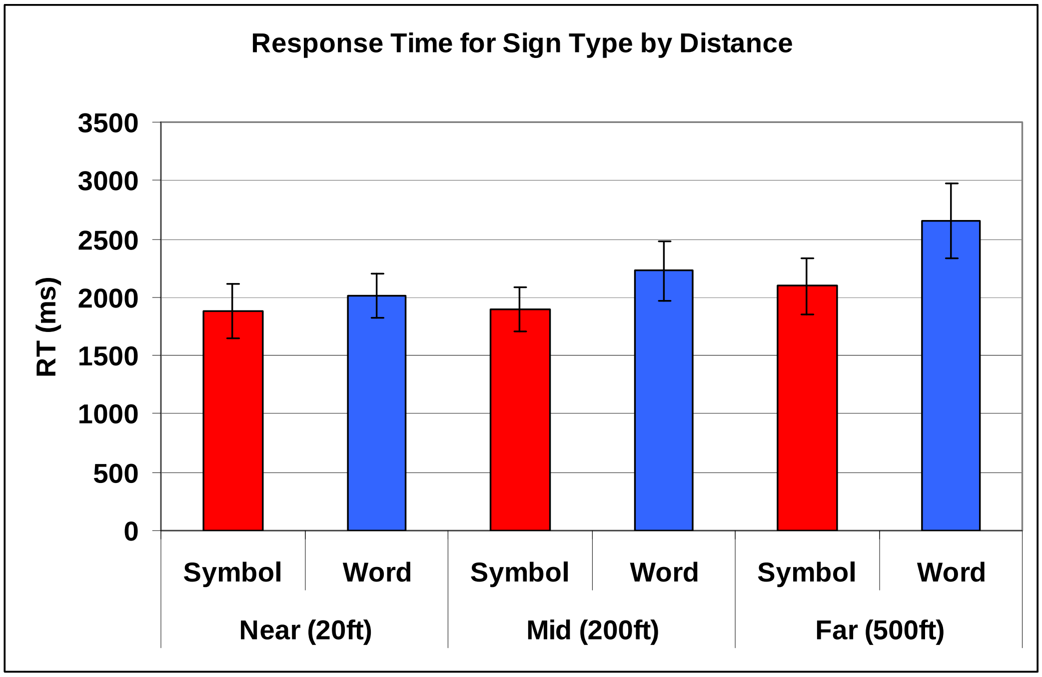


Figure 23. Comprehension Response time (ms) for sign type by distance. Error bars represent the 95% confidence interval for them means.

The interaction shows that the advantage in comprehension speed of responding to symbol signs over word signs increases with increasing distance. There is basically no significant advantage at 20 feet, but an increasing advantage for 200 and 500 feet respectively.

Pattern Matching Measures

These measures tapped the ability of drivers to indicate that a sign matched or mismatched.

Pattern Matching Accuracy

ANOVA showed that the only significant effects were age group, F (1, 61) = 7.36, p < 01, and distance (size of the sign), F (2, 61) = 116, p < .01, though there was a marginal (p < .08) interaction effect of sign type by distance. Older drivers were slightly less accurate (M=.86, SE = .01) then middle-aged and younger drivers who did not differ (M= .90, .91, respectively with SE = .01). Accuracy declined significantly at each distance from near (M=.95, SE = .006) to mid (M=.92, SE=.009) to far (M=.81, SE=.011). However, on the whole people show very high accuracy on pattern matching, higher than on comprehension accuracy, indicating that type of task has an impact on overall performance levels. Because of the varying levels of accuracy, we analyze speed of pattern matching only on correct trials.

Pattern matching speed

Similar to the results for comprehension speed, pattern matching speed depended significantly on distance, F (2, 122) = 62.2, p < 01, on sign type, F (1, 122) = 26.1, p < 01, and on the sign type by distance interaction, F (2,122) = 9.73, p < 01. There was a marginal effect of age, F (1, 61) = 3.06, p < 06, with younger drivers marginally faster than middle-aged and older drivers who did not differ from each other. The interaction is shown in Figure 24.

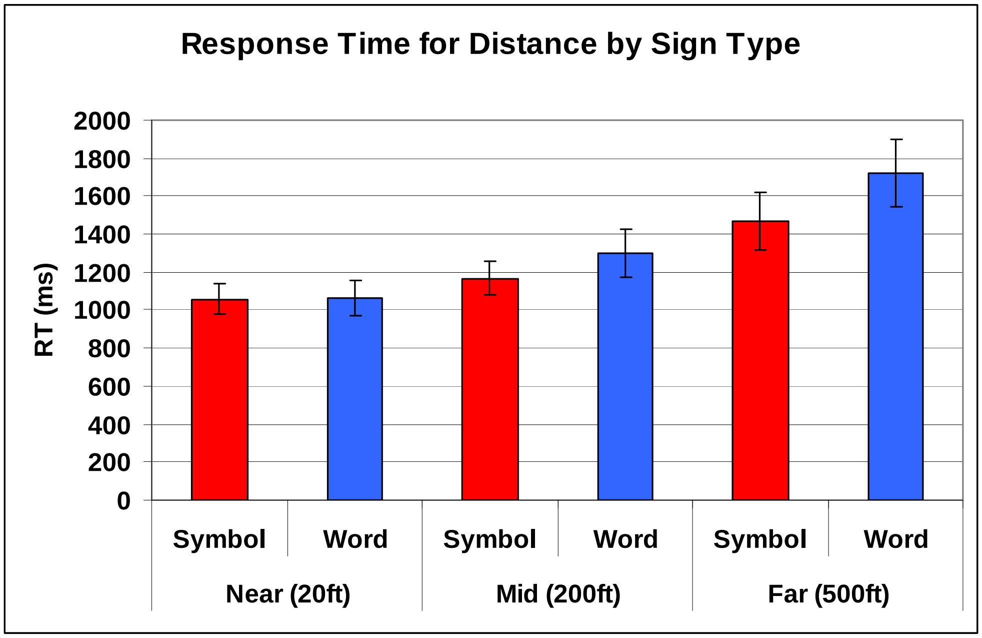


Figure 24. Matching Response time (ms) for sign type by distance. Error bars represent the 95% confidence interval for them means.

Similar to the case for comprehension time, the interaction shows no difference in speed of responding to symbol versus word signs at the near distance, but an increasing advantage for symbol signs over word signs at mid to far distances. Response speed generally slows with increasing distance (decreasing size of sign).

Study 3 Field Task.

Design

Task 3 field examined the viewing distances for signs with symbols versus words in a 3(age: younger/middle aged/ older) x 2(sign type: symbol/word) x 2(session: day/night) mixed design. Age was the between-subjects factor and sign type and condition were within-subjects factors.

Participants

A maximum of 16 younger (M = 25 yr), 18 middle-aged (M = 57yr) and 18 older (M = 71 yr) drivers participated across day and night sessions. Participants were recruited and screened in the same manner described in Study 1.

Materials and Procedure

The materials and procedure were identical to Study 1 Field except that there was no pattern matching task and different signs were used. The present set included 5 warning signs (36 x 36 inches) and 1 regulatory sign (24 x 30 inches). The signs included (see Figure 25): speed zone ahead, stop ahead, right lane ends, divided highway, hill ahead, and keep right. Each sign occurred once in the test set.

Participants were read a phrase conveying the semantic content of one of the signs and were instructed to drive forward at a comfortable speed until they could determine with 100% confidence whether or not the actual sign at the end of the track had the same meaning as the phrase they had been read. For half of the total trials, the phrase matched the actual sign, and on other half, the phrase did not match the sign. Once again, distance and speed were measured on each trial with a radar gun, with data recorded on a laptop computer.

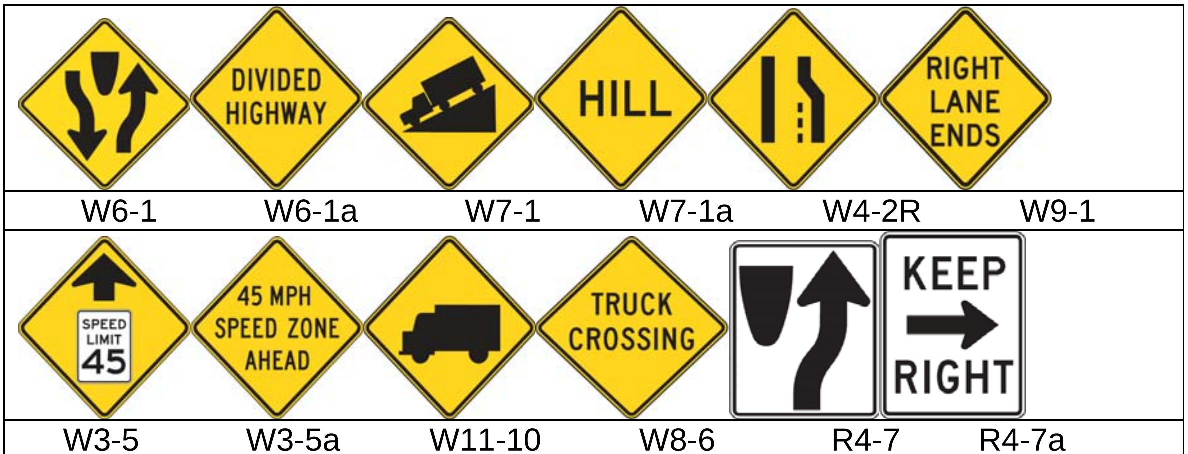


Figure 25. Signs used in the Task 3 field study.

Results

Field Study 3.

We had significant missing data due to corruption in a database file with complete data from 12 young, 18 middle-aged, and 16 older drivers during the day and 14, 15, and 16 respectively at night. The ANOVA examined the effects of age (young, middle, older), session (day, night), and sign type (symbol, word).

Accuracy

The only significant effect was session by age group: $F(1, 28) = 3.90$, $p < 05$. Figure 26 shows that accuracy did not vary for middle-aged drivers between day and night (86%, 90%). However, in the younger driver group, unexpectedly, accuracy was higher at night than during the day (86% compared to 78%) whereas in the older driver group it was higher during the day than the night (87% vs. 80%). Nonetheless, accuracy was reasonably high across conditions and age groups, suggesting that stopping distance scores could be interpreted reasonably reliably.

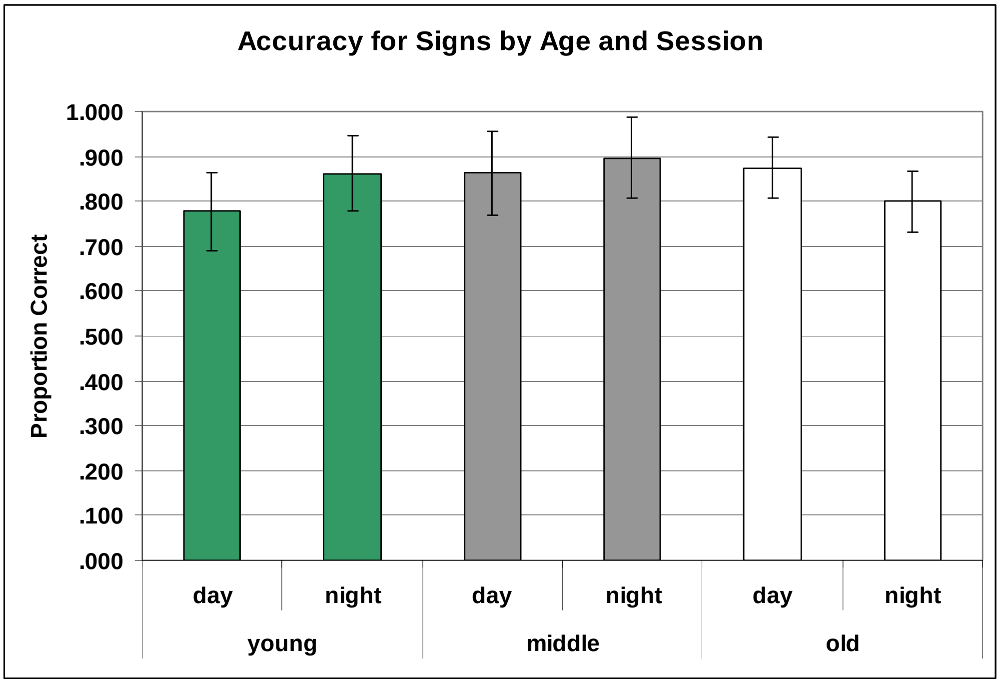


Figure 26. Mean proportion correct for signs (averaged across word, symbol types) by age group and by session (day, night). Error bars represent 95% confidence intervals for the means.

Stopping Distance

We had full data on 10 young, 14 middle-aged, and 14 older drivers. Age was not a significant factor ($F(2, 35) = 1.68$, $p > .2$, though power was quite low at .33 (one chance in three of detecting an effect of this size). Though there was a 100 foot stopping position difference between older drivers (547 feet), and the middle-aged (659 feet) and younger drivers (677 feet), it was not statistically significant given the high uncertainty associated with the estimate of the mean.

The significant factors were session (day, night), $F(1, 35) = 7.96$, $p < .01$ and sign type (symbol, word) ($F(1, 35) = 148$, $p < .01$. Stopping distances were considerably longer during the day ($M = 668$ feet, $SE = 36.5$) than at night ($M = 587$ feet, $SE = 33.7$) as expected. Symbol signs were much more legible ($M = 782$ feet, $SE = 40$) than word signs ($M = 473$ feet, $SE = 27.8$). Because power is so limited with this small sample size, we cannot be confident that small effects observed might not be significant in the general population, but the current significant effects are likely to be robust. Thus, there is a very clear advantage for the symbol signs over their word variants.

Comparisons between lab and field results.

To see to what extent lab performance maps to field performance we correlated average comprehension response time in the lab on symbols and words with average stopping distance in the field for that comprehension task, both during the day and at night for symbol and word warning signs.

Figure 27 shows the results for symbol signs in the lab and during the day in the field study, with an outlier present (top panel) and removed (bottom panel).

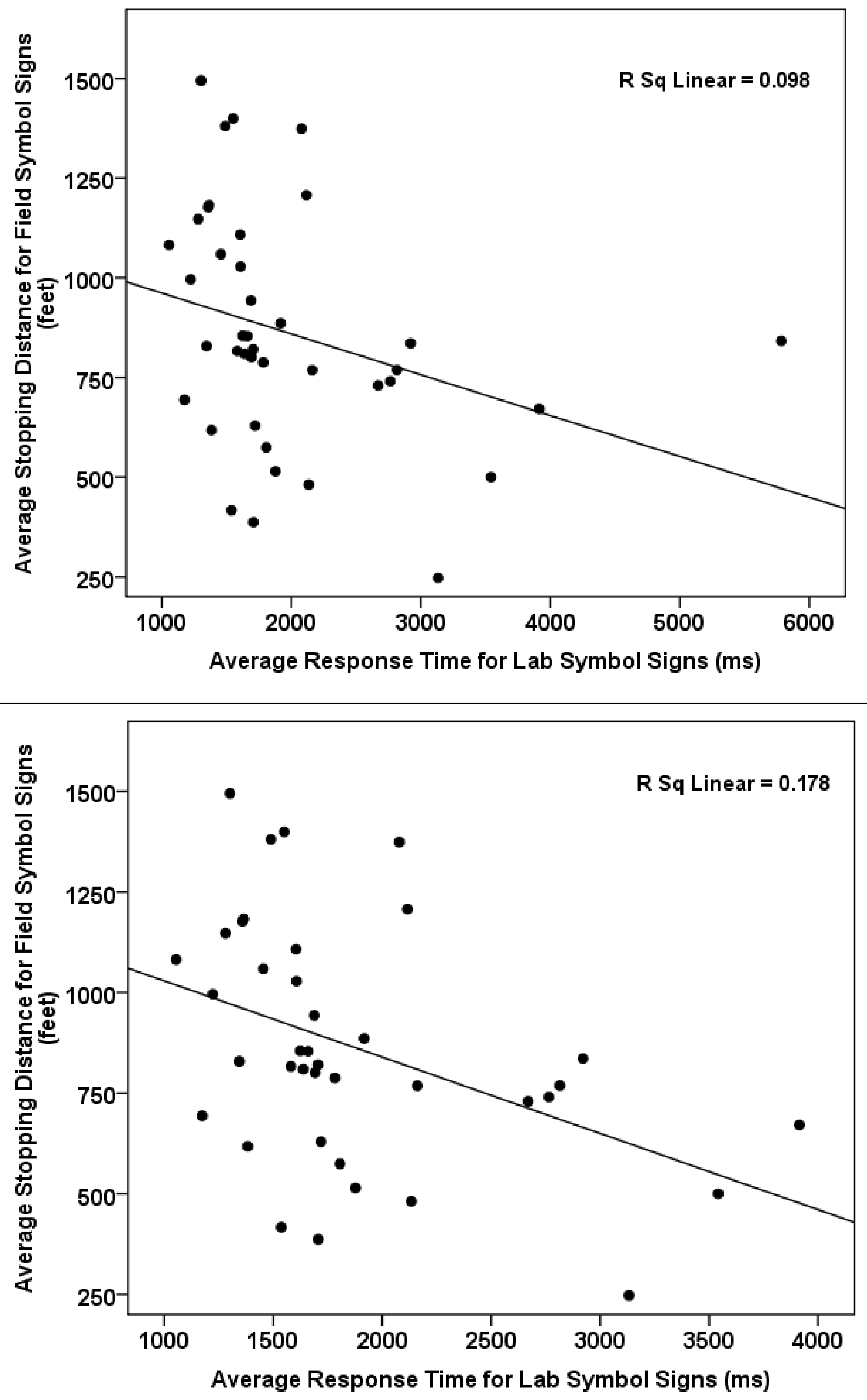


Figure 27. Correlation between Lab symbol sign response time (ms) with Field stopping distance on symbol signs (feet). Top panel contains an outlier (response > 5 s) and the bottom panel removes that person.

There is a weak relationship with the outlier case, but it strengthens considerably, accounting for 18% of the variance. Both represent statistically significant relationships between lab and field performance.

Figure 28 shows the relationship between lab word sign response time and field stopping distance for word signs.

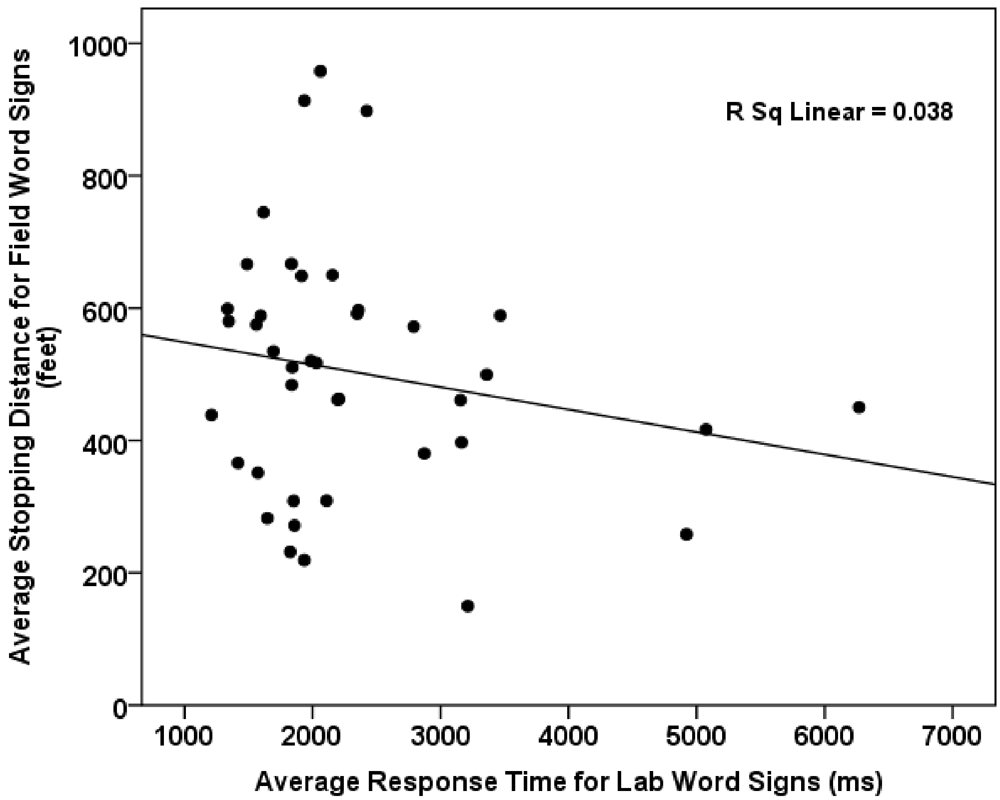


Figure 28. Correlation between Lab word sign response time (ms) with Field stopping distance on word signs (feet).

In this case there is no longer a significant relationship between lab and field performance. Removal of the outlier does not change the relationship much (r^2 =.04).

When analyses were conducted for the relationship between lab symbol performance and field night symbol performance, the relationship was not significant (r^2 = .001). The same was true for the relationship between lab word performance and field night word performance (r^2 = .017).

Thus, study 3 showed mixed support for predicting from lab performance to field performance. Thus, it is prudent to supplement lab studies with field studies when assessing sign legibility.

Study 3 Conclusions.

The field findings indicate that symbol signs are more quickly and easily understood compared to word signs. Thus, to optimize sign efficacy, symbol signs are to be preferred to word signs. It is also highly likely that for non-native English speakers and particularly for international visitors (tourists), symbol signs will be superior to word signs for warnings.

Although lab studies offer greater potential control over legibility variables, they need to be supplemented by field studies given the inconsistent relationships observed between performance in lab and field.

Caveat

We examined comprehension decisions with our studies. Whether drivers would be better served by the earlier recognition of the symbol compared to the word sign in terms of their driving response has not been evaluated (e.g., slowing, lane changes).

Chapter 5.

Overview of Studies 4-8.

Given that middle-aged drivers typically showed performance midway between younger and older drivers in the prior lab and field studies, we conducted the following studies in spring and early summer with just younger and older adults. The studies described below were carried out in a single session to minimize attrition.

Participants and General Procedure

Individual participation in Studies 4, 5, 6, 7, and 8 occurred in a single session. A maximum of 23 younger (M age = 25.3) and 23 older drivers (M age = 72.6) participated in the studies. The studies were completed in the following order: lab task, task 7, task 8, task 6, task 4. However, when couples were run the order was changed for the second participant to: task 7, task 8, task 6, task 4, lab. Occasionally overlapping of participants occurred and the order used was lab task, task 6, task 4, task 7, and task 8.

Several participants were unable to complete all of the studies because of humidity and intense temperatures that occasionally exceeded 100 degrees Fahrenheit, so sample sizes vary slightly across studies as seen in the results sections. Seven participants in total failed to complete all studies after having completed at least one of the initial studies.

Study 4: Word Messages on Portable Changeable Message Signs (PCMS)

Section 6F of the Manual on Uniform Traffic Control Devices (MUTCD) indicates that PCMS messages should be visible from 800 m (0.5 miles) under ideal day and night conditions. The goal for such signs is to convey messages about unexpected situations that will be readily understood by drivers and thus will allow them adequate time to react. As we have seen in prior studies reported above, older drivers showed poorer matching and comprehension skills for standard message signs than middle aged and older drivers. In this study, using the maximal distance possible at the TERL facility, 900 feet, we assessed how well a standard large PCMS message set could be perceived and comprehended by older compared to younger drivers. The expectation from MUTCD guidelines was that it should be comprehended at this distance.

Method.

Design

The study was a between subjects design with age the single factor with two levels: young and older drivers.

Materials

The portable message device used was the Precision Solar Controls Model SMC1000HE (see Figure 6). Data was collected using a Gateway 285-E laptop computer using the E-prime software package.

Eight pairs of standard steady (non-flashing, non-moving) messages were used for the experiment. These messages were: Minimum Speed 25 MPH, Minimum Speed 35 MPH, Keep Left, Keep Right, Reduced Speed 25 MPH, Reduced Speed 35 MPH, Road Work Next 2 miles, Road Work Next 5 Miles, Bridge Work 1000ft, Bridge Work 2000ft, Accident on Right, Accident on Left, Maximum Speed 25 MPH, Maximum Speed 55 MPH, Speed Limit 25 MPH, Speed Limit 55 MPH. Each message was displayed on a standard portable changeable message sign. See Figure 29. The sign face dimensions were 6'5" by 10'. The sign was raised to a height of 7 feet.

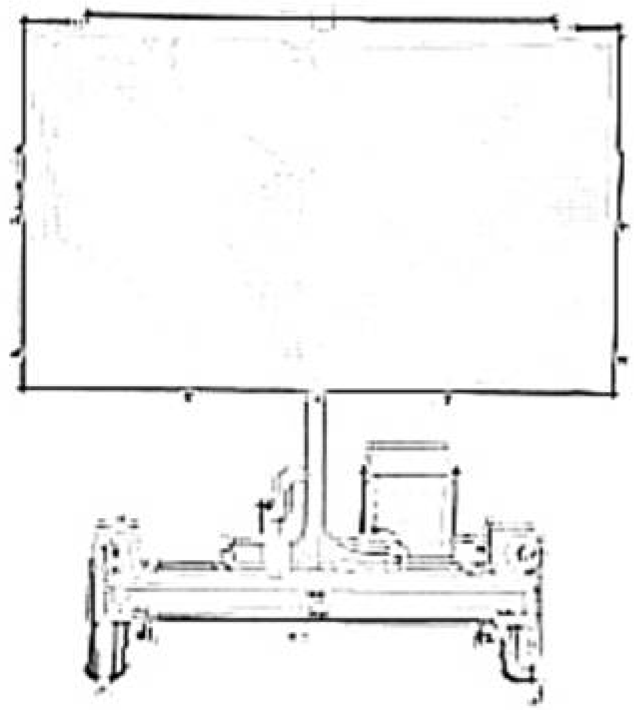


Figure 29. Illustration of the device used in Task 4.

Procedure

Drivers were asked to view the sign during the day only given that viewing distances are expected to be longer at night than during the day (contrast ratio of bright characters to dark backgrounds are enhanced when there is little ambient light on the PCMS display). Given the greater sensitivity of the matching procedure in earlier studies, we also used a match/mismatch judgment procedure for assessing sign comprehension. At the start of every trial the participant was situated 900 ft from the PCMS and stood with his/her back to the apparatus. The participant was presented with a message on the laptop computer and then instructed to turn around to face the PCMS. The participant was directed to walk forward at a comfortable pace until it was possible to determine if the message on the sign was the same as the message on the laptop. They indicated by saying "yes" if the message was the same, and "no" if it was different. Each participant completed eight trials and was presented with four matching messages and four mismatching messages in random order. One weakness of our procedure is that it introduced a memory component to the comprehension procedure, namely, participants needed to remember the content of the message for a second or so when they turned from the computer to the sign, although the memory load is not particularly high (2 – 5 words).

Results.

We have complete data from 23 younger adults and 19 older adults.

Accuracy

An ANOVA showed that there was a marginal effect of age group on accuracy (did the sign message match the verbal message), F (1,40) = 3.26, $p < .08$. Younger drivers were 96% accurate (SE = .025), and older drivers were 90% accurate (SE = .027). Thus, although there was a slight advantage for younger drivers, it was not statistically significant, though in the predicted direction. Another way to characterize the age difference is to assess what percent of each age group had perfect accuracy. We found that 17 of 23 younger adults had perfect accuracy but only 10 of 19 older adults. However, these values are not statistically significant when tested with a chi-square analysis.

Distance

Recall that people were asked to walk toward the sign until they could accurately determine if the verbal message matched the sign message. To retain as much data as possible, given the low number of trials and high accuracy (90% +), we analyzed both correct and incorrect trials. An ANOVA showed a significant effect of age group, F (1,40) = 11.2, $p < .01$. Younger adults made their decisions at 899.9 feet (SE = 5.64), whereas older adults moved closer to the sign, making their decisions at 871.9 feet (SE = 6.20). A more conservative analysis looking just at decisions on correct trials yielded the same findings.

Conclusions.

Despite having to move significantly closer to the sign by about 30 feet, older drivers proved slightly less accurate than younger drivers at comprehending PCMS messages. Whereas we do not know at what distance younger drivers would first comprehend the sign, given the limitations of the distances available at this test facility, it is clear that older drivers are not fully comprehending the messages within a half mile of the sign, as proposed by MUTCD guidelines. It is possible that comprehension processes for relatively lengthy text messages on a PCMS are simply not going to approach asymptote (100% accuracy) at any distance. However, using a very conservative procedure, having drivers walk toward the sign rather than drive toward it (which maximizes their opportunity to process part of the message at much lower velocity), it is evident that older adults need additional visual and cognitive support to reach higher comprehension levels at greater distances from the PCMS. This may require larger lettering, different lettering (change of font), different foreground and background lighting, or more efficient wording. An example of the latter would be putting the critical information first. For instance, changing the message from “Reduced Speed 25 MPH”, to “25 MPH REDUCED SPEED” might enable the driver who can only process part of the message to know what action to carry out. Also, our procedure tested the first phase of responding to a warning sign, namely, comprehension of the message, and we need further information about how quickly (at what distance from the hazard being signaled by the CMS) a driver will take appropriate action (e.g., slowing speed to 25 mph).

Caveat

Our field procedure required a memory component that may have disadvantaged older drivers.

Chapter 6.

Study 5: Lab Study of Comprehension of Countdown Pedestrian Signals.

One of the dilemmas facing a pedestrian about to enter an intersection is determining whether there is adequate time to cross safely. Even if a traffic signal indicates that it is permissible to enter the intersection (green light, walk signal), the pedestrian may be faced with a complex estimation process to determine whether the time remaining to cross is adequate in light of the length of the intersection, their own walking/running speed, congestion in the intersection, weather conditions, incline, etc. Botha et al. (2002) showed that San Jose pedestrians were relatively accurate at estimating the time to cross short intersections but had underestimates for longer ones. See Dunbar, Holland, and Maylor (2004) and Stollof, McGee, and Eccles (2007) for general reviews.

As mentioned earlier, one of the predictable consequences of aging is general slowing in physical and mental performance (Fisk et al., 2009). Large sample observational studies show that older pedestrians, those appearing to be over 65 years of age, walk more slowly, <1 m/s for 15^{th} percentile walkers, than do younger pedestrians, 1.25 m/s for 15^{th} percentile walkers (Knoblauch, Pietrucha, & Nitzburg, 1996). However, walking speeds are also a function of the walking environment (Finnis & Walton, 2008).

Countdown signals that display number of seconds left for crossing have been introduced to improve pedestrian decision making. Generally, pedestrian countdown signals are adequately understood by both younger and older pedestrians (Stollof, McGee, and Eccles, 2007), though about a third take too conservative an interpretation that the number is the time left before the light turns red rather than the majority interpretation that it is the time in seconds left to complete the crossing or reach a median. Nonetheless, comprehension is better for these displays than for the countdown don't walk ones (flashing hand).

However, observational studies have indicated that the introduction of countdown signals has produced mixed results in terms of pedestrian behavior. A Florida study by Huang and Zegeer (2000) showed that a countdown equipped intersection, matched with similar intersections that did not contain countdown signals, resulted in reducing the number of pedestrians who started running when the flashing "Don't Walk" signal occurred. The countdown-equipped intersection did not affect the number of people running out of time while crossing, and had a negative effect of reducing compliance with the "Walk" signal. However, other studies have shown different results, for instance, Schattler, Wakim, Datta, and McAvoy (2007) showed no increased risk taking by pedestrians.

Assuming that they are well calibrated about their own walking speed, that is, have adequate knowledge about their abilities, older pedestrians ought to be more cautious about entering intersections, and this can perhaps be seen in the longer time it takes them to step off the curb into the intersection when a light changes (about a half second longer than for younger pedestrians: Stollof, McGee, & Eccles, 2007).

A lab study by Singer and Lerner (2005) presented pedestrians with short (7 s) animations of a large countdown timer superimposed above a crossing scene with a pedestrian shown at the curb. Observers were asked to judge what action the pedestrian should take at the point in the display sequence when a beep was sounded.

The study also tested standard traffic signals ("don't walk" signal) and enhanced signals where the countdown was displayed without the "don't walk" signal. Understanding of the correct action to take was best with the countdown only display. In their second study they observed crossing behavior at intersections with different displays (standard don't walk, countdown plus don't walk, experimental countdown only), showing that pedestrian countdown displays elicited somewhat riskier crossing behavior (people crossing late in the walk interval) though they tended to arrive safely at the opposite side, suggesting that they hurried across knowing the minimal time for crossing. Intersections also incorporate a 5 s interval from the end of the pedestrian crossing interval (steady don't walk signal phase) until vehicles receive a green light so there is a buffer period for pedestrians.

In this experimental study, we assessed the decision making processes of younger and older pedestrians who were presented with pictures of different intersections equipped with countdown displays and asked them to make a decision about whether to cross or not to cross the intersection. In the next (field-experimental) study, we used existing intersection-like parts of the track at TERL to compare lab-based decision-making with decision-making using a real intersection and countdown signal system.

In our experimental study we contrasted decision making in the normal pedestrian crossing situation of a countdown display with that for a control condition, presentation of the same information in analog format using a dynamic pie display. The pie display showed full time as a filled red circle and blacked out an appropriate proportion of the circle with passing time until it became unfilled. Although there are no pedestrian displays corresponding to this format, this analog control condition enables us to explore pedestrian decision making processes similar to those with a flashing hand but with a closer analog to digital count information. (Similar analog displays are commonly used to indicate time remaining for an operation with computer displays.)

Method.

Design

Study 5 assessed participants' willingness to cross intersections under a variety of conditions in a simulated environment. Age group was a between-subjects factor, and timer and intersection length were within-subject factors in the 2(age group: younger/older) x 2(timer: countdown/dynamic pie) x 3(intersection length: short/intermediate/long) design.

Materials and Procedure

Participants were seated in front of a Gateway M285-E laptop computer equipped with MediaLab software (http://www.empirisoft.com/). Each trial began with a 3-second video of an intersection with a timer displayed in the top right corner. We chose 3 s in part because it is slightly above the suggested 2.5 s perception-response time for processing signals by drivers and might be similar to a short glance at the signal by a pedestrian. A black screen immediately followed the disappearance of the video where the participant had to click the mouse button once to advance to the question portion. The participant then had to use either the laptop pad or the USB mouse to click on the answers to the questions. The first question was whether or not they would cross and the answer choices were "yes" or "no". The second question was how safe they feel the intersection was and the answer choices were not safe at all, slightly unsafe, neutral, slightly safe, or very safe. The third question was how fast they would have to walk if they had decided to cross and the answer choices were normal pace, slightly faster pace than normal, or they would have to rush. After all three questions were answered, the participant clicked to continue onto the next video presentation.

There were 30 total videos with the first half containing a regular countdown timer with flashing hand and time value, and the second half containing a dynamic red, pie-shaped timer that showed elapsing time with an increasing black gap in the circle opening from the east part of the circle. In each half, there were 10 short intersections,10 medium intersections, and 10 long intersections. The intersections used were sampled from the Leon County area and were analyzed and grouped by distribution so that short intersections are between 40 feet and 57 feet across, medium intersections are 58 feet to 87 feet, and large intersections are 88 feet to 189 feet. The countdown timers were recorded from the regulation sized pedestrian device used in Study 6. The countdown timer video was then sampled for the appropriate time frame for each video shown to participants. The pie-shaped timer was acquired using a java-based script for displaying pie-shaped timers. Static examples from the clips can be seen in Figure 30. The same intersection video order was used for all participants.

After the participant finished with the video portion, they were then asked about what factors influenced their decisions to cross or not cross. The participants used the laptop keyboard to type one factor and then pressed enter to go to a new screen in which they could add another factor. There were 5 screens for entering these factors and the participants could type N/A to skip an answer if they could not think of anything else.

Next, the participants were asked to type in how often they usually cross intersections similar to the ones shown in the videos. This question was open-ended and produced both quantitative (e.g. "3 times per week") and qualitative (e.g. "rarely") answers.

Finally, the participants were asked if there were any other factors that they would consider in crossing the street that the lab exercise did not cover. Participants typed in one factor and pressed enter to go to the next, for a total of 3 possible factors. They also had the option to type N/A to skip an answer if they could not think of anything else.

Figure 30. Task 5 static pictures of pedestrian countdown display (left) and countdown pie display (right)

Results.

Comfort with crossing decision

An ANOVA on perceived safety (1 = not safe, to 5 = very safe) with crossing showed significant effects of age group, $F(1, 44) = 12.9$, $p < .01$, as well as for length of intersection (short, medium, long), $F(2, 88) = 11.4$, $p < .01$, and an interaction of type of countdown display (pie, numeric) by length of intersection, $F(2, 88) = 11.8$, $p < .01$. There was a marginal interaction of type of countdown by age, $F(1, 44) = 3.96$, $p < .06$. The age group effect showed younger pedestrians feeling more safe (M = 3.55, SE = .153) than older ones (M = 2.77, SE = .153), which as will be seen in study 6, corresponds well to the reality of making a safe traversal.

The interaction of length by type of countdown is shown in Figure 31. It indicates that the standard countdown timer yielded slightly lower safety feelings for short than medium or long intersections, whereas the pie countdown timer yielded safety ratings that declined from relatively safe for short and medium intersections to relatively unsafe for longer ones. That is, the pie presentation left people better *calibrated* about the perceived safety of crossing the intersection, assuming that the length of time in an intersection is a risk factor. The marginal interaction of age with type of countdown display showed that younger pedestrians did not vary in their feelings of safety across display types (M =1.2, SE = .056 for numeric; M = 1.2, SE = .047 for pie), whereas the older adults trended in the direction of feeling safer with the numeric (M = 1.3, SE = .056) than the pie (M = 1.2, SE = .047) display.

We obtained the same outcome (ANOVA results) when using a 1 (comfortable pace) – 3 (need to rush) rating scale about how quickly the pedestrian would need to cross the intersection.

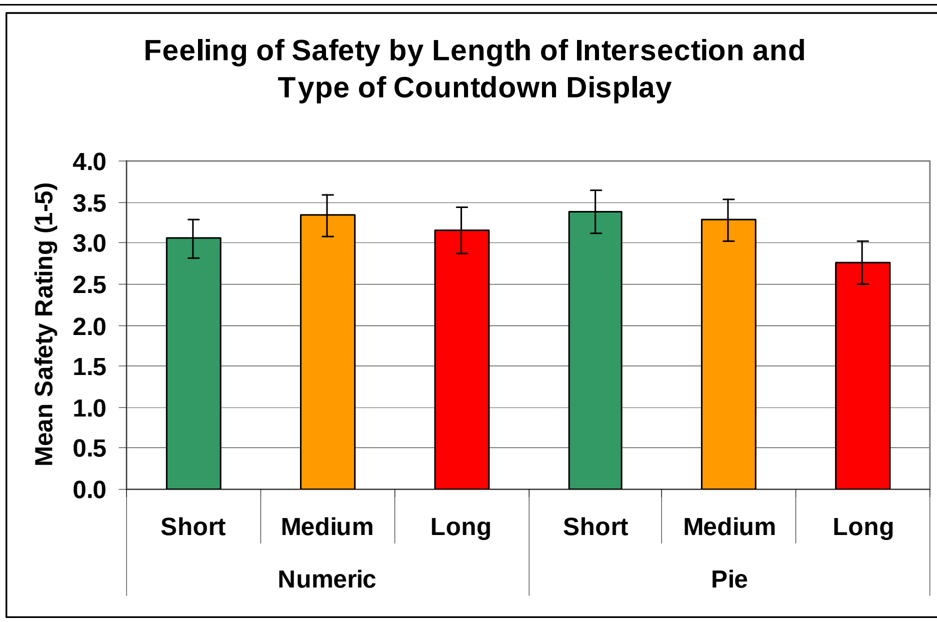


Figure 31. Mean rating of safety (1= not safe to 5 = very safe) by length of intersection and type of countdown display.

Decision to cross

We recoded the decision to cross as 0 for no and 1 for yes, so means will represent the proportion of positive cross decisions. ANOVA showed significant effects of type of countdown display, $F(1, 44) = 9.09$, $p < .01$, length of intersection, $F(2, 88) = 3.40$, $p < .05$, but no effect of age group, $p > .25$. A significant interaction was seen for countdown type by length, F (2, 88) = 15.2, p < .01. The interaction, which is similar to that for comfort rating, can be seen in Figure 32. Here too, the standard numeric countdown display leads to a higher probability of deciding to cross as intersection lengthens (with surprising conservatism at the shortest interaction), whereas the pie display generate a higher probability of positive crossing decisions for short and medium intersections but a lower one for long intersections.

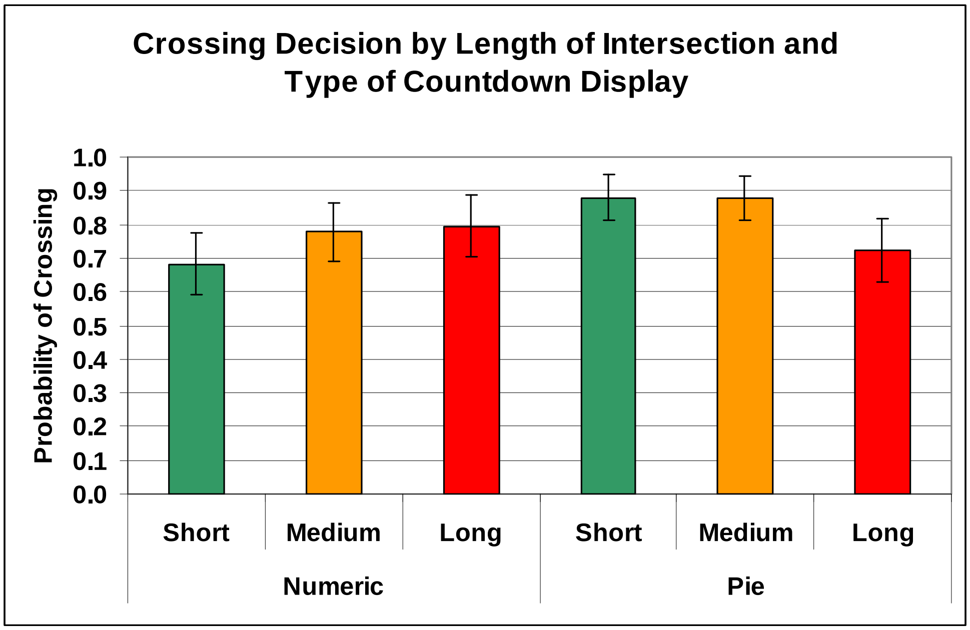


Figure 32. Probability of a positive crossing decision by length of intersection and type of countdown display.

Conclusions.

Older pedestrians, who are at greater risk of injury in intersection crashes, generally felt less safe about crossing an intersection than younger pedestrians. This discomfort is probably adaptive given their slower walking speed which puts them at greater risk of being struck by an inattentive driver. The pie countdown display yielded a greater proportion of positive crossing decisions for the short and medium intersections, and induced greater caution for the long intersections, a desirable result. The numeric countdown display generated somewhat greater caution overall, particularly for short intersections. Prior literature shows that countdown timers (compared to flashing hand displays) promote slightly risky behavior where pedestrians tend to rush across the intersection, though they typically arrive at the other side before the light changes. It would be worthwhile to consider additional research with a pie countdown display, particularly manipulating the time remaining and doing so in a field setting.

Caveats

As is the case with lab studies generally, decision making while watching short film clips may not translate well to the crossing decision at a real intersection, particularly one with other pedestrians present, at night, or with adverse weather conditions. Also, an animated pie timer is a novel display for all pedestrians so comparisons with a countdown timer, now fairly widely deployed, may reflect novelty rather than represent true motivational differences.

Chapter 7.

Study 6: Field Study of Comprehension of Countdown Pedestrian Signals.

Although an experimental investigation of decision making in a lab setting is helpful, it is also useful to conduct field experiments where we can manipulate the state of the signal in more realistic settings. We expected that older adults would generally be more conservative about crossing (given their slower walking speed than young adults). Also, we used this field study as an opportunity to assess the legibility of the countdown signal, given the unexpectedly weak performance of older motorists with PCMS messages in study 4.

Method

Design

Study 6 estimated the legibility distance of countdown timers on intersection pedestrian signals and assessed participants willingness to cross intersections as a function of time in a 2 (age group: younger/older) x 3 (duration: 29 seconds/19 seconds/1 second) x 2 (having viewed a walk symbol for five seconds/having not viewed a walk symbol).

Materials

This task implemented a standard 16” pedestrian countdown device (see Figure 33). This device was seated atop an 8’ pole. The countdown device was controlled using a standard traffic signal control device. Countdowns were manually triggered using a test switch.

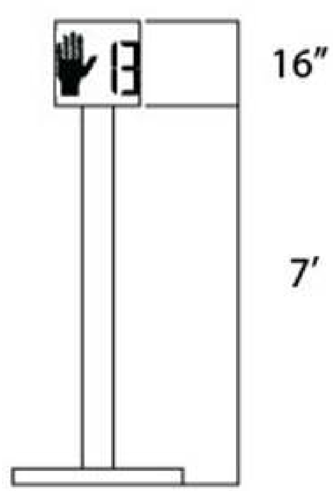


Figure 33. Diagram of the device used for Task 6 field study.

Procedure

We adopted MUTCD guidelines for setting the countdown period, assuming a walking speed of 1.2 m/s (4 ft/s) for allowing a pedestrian adequate time to reach the far side of the intersection with a start at the beginning of the cycle (e.g., the walk signal). We calculated for each of the distances (50, 75, 115 feet) pedestrian clearance times as if they were given the full clearance time. For example, 50 feet at 4 ft per second equals 12.5 seconds which we rounded up to 13 seconds. We used the 13 seconds as the pedestrian clearance/countdown time to give participants the maximum chance to finish crossing the road within MUTCD guidelines.

Individual participants were situated 400 feet from the pedestrian signal and stood with their backs facing the signal. The countdown timer was set to sixty seconds and was initiated as participants were directed to turn around to face the signal, and begin walking forward. They were instructed to begin reading the descending digits as soon as they were legible. Ground markers were used to record the participant's legibility distance.

Next, participants were presented with randomized trials in which they would walk a simulated intersection 50 feet, 75 feet, or 115 feet long. Each intersection length had a corresponding minimum regulation length pedestrian crossing time of 13 seconds, 19 seconds, and 29 seconds, respectively. Participants were instructed to turn their back to the device and were instructed to turn around either when the walk symbol appeared and remained for five seconds before countdown, or after the walk symbol cleared and countdown had begun. In either case, the countdown starting point was identical. Participants completed six trials with two trials (having viewed a walk symbol vs having not viewed a walk symbol) at each intersection. The flash don't walk hand symbol (FDW) occurred during the countdown cycle and remained steady at the end of the cycle.

Additionally, participants were instructed to finish crossing once they began regardless of the time remaining on the countdown. It was also noted if a participant was not successful in crossing defined as the countdown timer reaching zero before they reached the other side of the intersection. Two participants traveled less than 10 feet before turning back reporting that they would not cross. These were both recorded (conservatively) as would not cross decisions.

Results.

Legibility for Pedestrian signals.

We have complete information from 20 young and 18 older pedestrians.

Legibility accuracy

We assessed proportion correct reports for reading the countdown from the distance at which the pedestrian thought they could report the information accurately. An ANOVA looking at age group showed no significant effect, $F(1, 36) = 2.37$, $p > .10$. Average accuracy (proportion trials correct) was .944 (SE = .036).

Legibility distance

An ANOVA using both correct and incorrect trials comparing younger and older pedestrians showed a significant effect of age on distance, $F(1, 36) = 10.8$, $p < .01$. A similar effect is obtained when just correct trials are included in the ANOVA. Younger pedestrians were able to report the countdown value from a distance of 375 feet (SE = 51.9), whereas older pedestrians moved closer, to 307.5 feet (SE = 73.7). However, it is evident that neither group would be expected to have problems with perceiving the information about the countdown for normal intersections which are much shorter than 300 feet.

Decision making for pedestrian signals.

Crossing decision

We examined age differences using chi-square analyses of the choice to cross or not to cross as a function of the intersection length (short, medium, long) and the timer state (partial: FDW symbol present / full time allocated: no FDW symbol present). As seen in Table 3, there were significant age differences for the case of a long intersection (115 feet) with full crossing time available, chi-square (1) = 4.55, $p < .05$. Older pedestrians were more conservative.

Table 3. Decision to cross by age group for the long intersection full-crossing time condition.

Crossed		Long Intersection with Full Crossing Time		
		yes	no	Total
Age Group	Younger	21	1	22
	Older	15	6	21
	Total	36	7	43

We observed the same effect for the long intersection with a partial countdown, as seen in Table 4. Here older adults were even more conservative, chi-square (1) = 11.3, $p < .01$.

Table 4. Decision to cross by age group for the long intersection partial-crossing time condition.

Crossed		Long Intersection with Partial Crossing Time		
		yes	no	Total
Age Group	Younger	20	2	22
	Older	9	12	21
	Total	29	14	43

For all other conditions (medium, short intersections with full or partial countdowns), there were no significant differences between younger and older pedestrians, with most pedestrians choosing to cross.

Crossing Success

For the crossing decision, experimenters categorized the cross attempt as “successful” and as “time ran out” (or didn’t try to cross). We used chi-square analyses to assess any differences between younger and older pedestrians depending on the length of the intersection (long, medium, short) and the state of the countdown (partial, full). We restricted analysis to cases of attempted crossings. Recall that older adults were generally less willing to attempt to cross, as seen above, so any differences are likely to be practically significant. Table 5 shows that for those older adults choosing to cross, compared to younger pedestrians they were less successful on the long intersection with full time, chi-square (1) = 6.3, p < .02.

Table 5. Crossing success by age group for the long intersection full-crossing time condition.

Success		Long Intersection with Full Crossing Time		
		yes	Time ran out	Total
Age Group	Younger	21	0	21
	Older	11	4	15
	Total	32	4	36

However, for the long intersection with partial crossing time there were no age differences, chi-square (1) = 2.19, p > .10

For the medium intersection with full crossing time, again there was a significant age difference in success, chi-square (1) = 4.02, p < .05, with older adults less successful as seen in Table 6.

Table 6. Crossing success by age group for the medium intersection full-crossing time condition.

Success		Medium Intersection with Full Crossing Time		
		yes	Time ran out	Total
Age Group	Younger	21	0	21
	Older	14	3	17
	Total	35	3	38

A similar pattern was evident for the medium intersection with partial time, where again older pedestrians were more likely to run out of time on the crossing, chi-square (1) = 6.18, p < .02, as seen in Table 7.

Table 7. Crossing success by age group for the medium intersection partial-crossing time condition.

Success		Medium Intersection with Full Crossing Time		
		yes	Time ran out	Total
Age Group	Younger	19	0	19
	Older	10	4	14
	Total	29	4	33

For the short intersections there were no significant differences between age groups in successful crossings for either full or partial countdown timings.

Walking speed

Although we estimated walking speeds for our sample, there were too many missing data points (refusals to cross, missing data) to generate sound estimates for an ANOVA that could provide information about the influence of age, intersection length and available crossing time.

Conclusions.

Standard countdown signals are highly legible at distances greater than most intersections for both younger and older pedestrians so current guidelines for their construction are adequate. Whether they are likely to be good enough for those with low vision is not known given that our older and younger groups were quite healthy.

Older pedestrians, despite being more conservative about deciding to cross an intersection were more likely to fail to cross in time than younger pedestrians for longer and medium-length intersections. Their conservatism about entering an intersection is well founded.

Caveats

Our intersections were not true replicas of real intersections and there was no significant traffic on the track so choices about crossing did not carry the same consequences. As well, the study was carried out in high heat and humidity, a time when people are less likely to be walking. Thus, our outcomes, while informative and consistent with observational studies, need further replication and extension.

Chapter 8.

Study 7: Analysis of pedestrian conflicts with left and right turning traffic

The goal in this study was to assess whether younger and older drivers are equally sensitive to a pedestrian encroaching on an intersection. There is evidence that velocity of the vehicle prior to approaching an intersection is a strong predictor of willingness to yield to a pedestrian entering an intersection (Geruschat & Hassan, 2005). There is also evidence that drivers over age 60+ are less likely to commit crosswalk violations (Kim, Brunner, & Yamashita, 2008). A number of measures have been used to examine when a driver detects a pedestrian (see Langham & Moberly, 2003) ranging from laboratory response time to pictures to field studies asking the driver to directly report when they notice the pedestrian by talking or by pressing a touchpad. These direct measures have the disadvantage of requiring the driver to look for pedestrians and are not that representative of normal driving. In this study we used an indirect procedure. We followed the progress of a vehicle with a radar gun to examine the velocity profiles for left and right turn situations by older and younger drivers as they approached, entered, and departed from intersections that were, or were not encroached on by a pedestrian.

Method.

Design

Study 7 assessed a driver's reaction to the presence (or absence) of a pedestrian encroaching on the road in the vehicle's turn path in a 2 (young/older) x 2 (pedestrian/no pedestrian) x 2 (left/right turn) design.

Participants.

The same younger and older drivers from the prior studies (studies 4 -6) participated in this one.

Materials

The Stalker Radar ATS was connected to a Dell Inspiron 700m laptop to record vehicle velocity profiles. For a driving track, the participant used the short track (see Appendix B) on the TERL facility property.

Procedure

The participant and experimenter entered the participant's vehicle and the experimenter instructed the participant to drive around the track several times so that measurements (unspecified to the participant) could be taken. The participant was instructed to turn left at the intersection during two laps and right at the intersection during another two laps. The traffic signal was disabled so that there was no need to stop the vehicle during these turns. Left hand turns came first, then right hand. Presence or absence of a pedestrian at the curb was randomized. During the turns, a confederate wearing normal street clothing waited on the curb of the street that the vehicle was turning into (see Appendix A). The confederate waited on the curb opposite the lane that the vehicle would occupy upon completion of the turn (outside curb for the right turn; inside curb for the left turn). The pedestrian was non-visible for one trial, and the next they were stepping out. During one left turn and one right turn, the confederate took several steps into the lane as the vehicle crossed into the intersection. Speed of the vehicle during the turns was measured with the Stalker Radar ATS.

Results.

We have complete data for 23 younger and 17 older drivers. The main variables we examined were start speed (velocity of the car when radar measurements were started), end speed (velocity at the end of radar measurements as drivers accelerated out of the completed turn), initial deceleration speed (velocity as driver started braking), final deceleration speed (lowest speed reached during a turn, which may be 0 if stopped for a pedestrian), and deceleration slope (an index of how quickly the driver decelerated during the turn). We averaged these variables across pedestrian present or absent and type of turn, given the general rule that there should be about 10 cases per variable for factor analysis, and performed a factor analysis to assess if they provided independent information. A single factor accounted for 63% of the variance, with factor loadings shown in Table 8. The results are consistent with the view that drivers have stereotypical driving speeds when entering, navigating, and leaving intersections during left and right turns.

Table 8. Factor solution for Radar Measures

	Component
	1
StartspeedAvg	.962
EndspeedAvg	.879
FinalDecentSpeedAvg	.283
SlopeAvg	-.656
DecentInitialAvg	.963
Extraction Method: Principal Component Analysis.	

Thus, these measures reflect a common underlying factor. We present results for those measures showing main effects or interactions that would not be reflected in the factor analysis which collapsed across pedestrian present versus absent and for type of turn.

Start speed

The initial radar measurements were taken as drivers approached the turn on the track, so may not reflect the influence of presence of a pedestrian if they were not monitoring far into the turn. ANOVA showed significant effects of age, turn ($p < .01$) and an interaction of age by turn, $F(1, 38) = 10.2$, $p < .01$, that can be seen in Figure 34. Presence or absence of a pedestrian had no significant effect and did not interact with other variables. The age by turn interaction indicates that younger drivers did not vary in initial speed for type of turn, but older drivers did, approaching more slowly on upcoming right turns more than on left ones. Right turns are somewhat more difficult than left ones because of the shorter turning radius available to the driver to complete the turn (assuming that they stay in their lane). Generally speed is low, given the driver's unfamiliarity with the track, and the relatively short distance available for driving on the track.

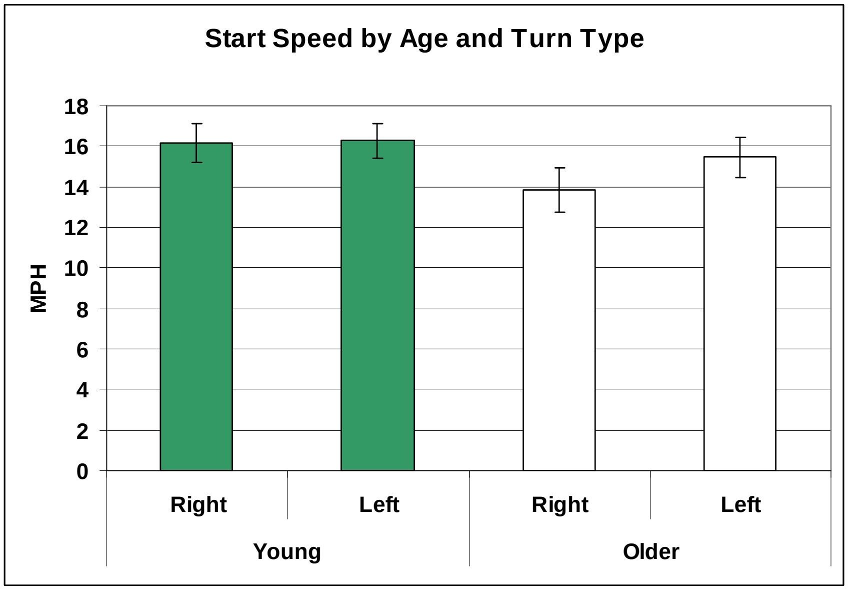


Figure 34. Speed (mean mph) entering before the turn by age and turn type. Error bars represent 95% confidence intervals for the means.

End speed

ANOVA conducted on the velocity at the end of the turn (past the pedestrian) showed only significant main effects of type of turn, F (1, 38) = 96.8, p < .01, and presence/absence of a pedestrian, F (1,38) = 11.3, p < .01, and a marginal effect of age (p < .06). Speed was higher coming out of right turns (M = 16.8, SE = .366) than left turns (M = 14.2, SE = .377). Speed was lower when there had been a pedestrian present (M = 15.1, SE = .351) than when no pedestrian was present (M = 15.9, SE = .389). The marginal age effect showed that older drivers were slower (M = 14.8, SE = .528) than younger drivers (M = 16.2, SE = .454). Thus, drivers of all ages showed sensitivity to the presence of a pedestrian on completion of the turn.

Initial deceleration speed

ANOVA on the speed traveled just before braking/deceleration revealed the same pattern as for start speed: significant main effects of age, turn (p < .01) and also a significant age by turn interaction that is nearly identical to that seen for start speed, and is shown in Figure 35.

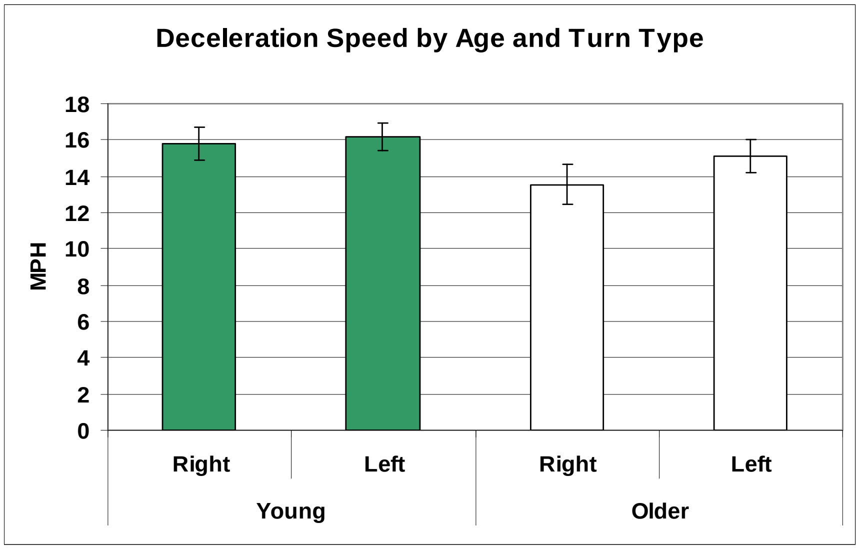


Figure 35. Speed (mean mph) before deceleration/ braking by age and turn type. Error bars represent 95% confidence intervals for the means.

Here too older drivers show more caution before right turns than do younger drivers.

Final deceleration speed

An ANOVA on the final (lowest) speed during the turn showed effects of age, $F(1, 38) = 4.55$, $p < .05$, pedestrian, $F(1, 38) = 68.5$, $p < .01$, turn, $F(1, 38) = 18.0$, $p < .01$, and an interaction of pedestrian by turn, $F(1, 38) = 19.7$, $p < .01$, with a trend toward a pedestrian by turn by age interaction ($p < .06$). Older drivers reached a lower speed ($M = 1.50$, $SE = .100$) than did younger ones ($M = 1.78$, $SE = .086$). Final speed was lower when a pedestrian had been present ($M = 1.13$, $SE = .115$) than when none was present ($M = 2.14$, $SE = .053$). Final speed was lower for a right turn ($M = 1.40$, $SE = .082$) than for a left turn ($M = 1.87$, $SE = .088$). The pedestrian by turn interaction is shown in Figure 36. It indicates that when a pedestrian is present right turns yielded greater slowing than left turns. Without a pedestrian present there was no reason to slow to an extreme extent in the intersection and the turn type had no effect.

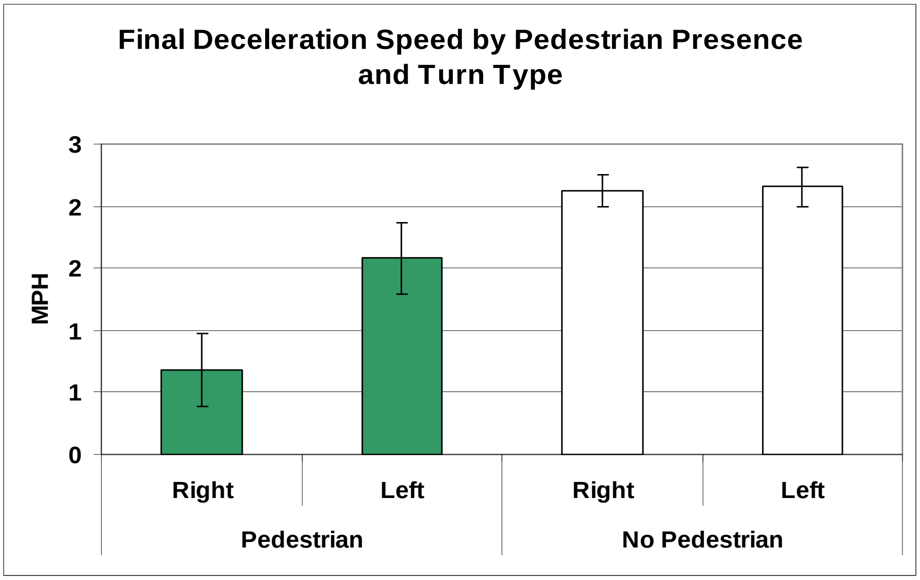


Figure 36. Final deceleration speed (mean mph) by pedestrian and turn type. Error bars represent 95% confidence intervals for the means.

Slope of deceleration

This variable captures how quickly the driver decelerated during the turn. ANOVA showed significant effects of type of turn, F (1, 38) = 9.75, p < .01, and of presence of pedestrian, F (1, 38) = 90.9, p < .01. Deceleration was greater for right turns (M = -.221, SE = .008) than left turns (M = -.202, SE = .010). Slowing was also greater for pedestrian present (M = -.257, SE = .011) than for pedestrian absent (M = -.166, SE = .009).

Conclusions.

Drivers of all ages showed some sensitivity to the presence of a pedestrian entering the lane opposite that of the driver at an intersection. However, there was little difference between older and younger drivers in response to the presence of a pedestrian. Older drivers drove more slowly on the track than did younger ones, and showed greater sensitivity to whether the turn they were to make was a right or left turn. Thus, for the case of a pedestrian wearing normal street clothes entering an intersection there is little difference in the behavior of a younger or older driver. Older drivers tend to drive more slowly on average, particularly for right turns.

Caveats

Most pedestrian crashes occur at night (Langham & Moberly, 2003). Older drivers have much poorer night vision than younger drivers and hence detect pedestrians at much reduced distances, particularly those wearing dark clothing (Wood, Tyrrell & Carberry, 2005). Further field research is warranted using our indirect measures in night conditions with pedestrians varying in conspicuity.

Chapter 9.

Study 8: Comparison of sign legibility under adverse visibility conditions

Field studies and experiments on sign perception are generally conducted under ideal weather conditions in an attempt to hold weather constant so that other variables can be properly assessed. This approach makes sense, given that most traffic crashes occur during the day in clear weather conditions (Evans, 2004). There is evidence that older drivers intentionally modify their driving habits to avoid difficult driving environments (Hakamies-Blomqvist & Wahlstrom, 1998). Examples are that they prefer to avoid driving during rush hour traffic, at night, on freeways, and in poor weather conditions (rain, fog, snow). A retrospective interview analysis of traffic crashes for drivers in British Columbia, Canada showed that many occurred when older drivers were delayed so that their return trip took place at night or when they were caught unexpectedly in changing weather conditions (Rothe, 1990). Florida experiences very heavy rain conditions, particularly in summer season when intense heat and humidity generates convection-based thunderstorms. It is during such adverse weather that older drivers probably need the greatest environmental support (Morrow & Rogers, 2008) for driving, including high conspicuity of traffic signs and signals. In this study, we simulated heavy rainfall on the vehicle windshield and assessed sign legibility as a function of sign type (symbol, word) and driver age (young, older).

Method.

Design

Study 8 was designed to assess the driver's ability to see warning signs in heavy rain in a 2 (age: young/older) x 2 (sign type: symbol/word).

Participants

The same young and older drivers from the prior studies (studies 4 -7) participated in this field experiment.

Materials

For this study, a rain-mimicking device was constructed (see Figure 37). The 10'x10' frame of this device was constructed of wood with PVC piping forming the plumbing atop wheels for mobility. Two rows of four rain-like sprinklers were used to form the 'rain'. Water was provided from underground plumbing on TERL property.

For the mobility of signs, a heavy-duty garden wagon was outfitted with a signpost to allow for the mounting of signs (see Figure 38). The overall height of this device was nine feet tall. The device could not be regulation sign height due to top-heavy tipping. Much like study one, signs were mounted together in a dual sign head for easy changing of signs. The signs used for this study were (see Figure 39): hill word (W7-1a), right lane ends word (W9-1), divided highway word (W6-1a), yield ahead symbol (W3-2), speed zone symbol (W3-5), stop ahead symbol (W3-1).

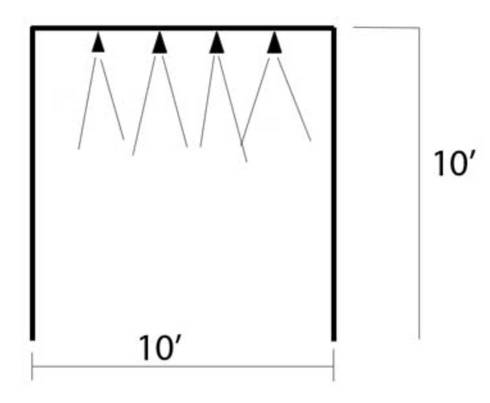


Figure 37. Diagram of the weather mimicking device for Task 8.

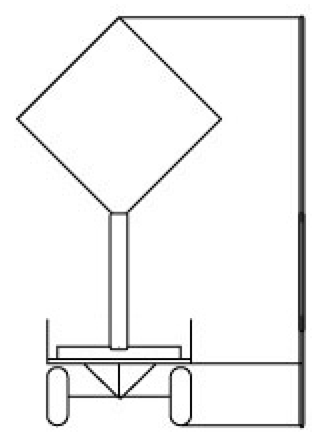


Figure 38. Diagram of the sign mobility device.

W7-1a	W9-1	W6-1a	W3-2	W3-5	W3-1

Figure 39. Signs used in Task 8.

Procedure

Participants parked their vehicles at a specified point on the track (maybe show in diagram) and remained inside with the experimenter for the duration of the task. The rain simulation apparatus was moved into place with the sprinklers adjusted to maximize the area of spray on the windshield. The sprinklers were activated and the participants were instructed not to turn on their wipers until directed to do so.

Each trial began when experimenters on the opposite side of the track affixed a new sign to the cart. As with other experiments, the experimenter presented participants with a verbal description of the sign stimulus on the laptop monitor (as well as information about where to expect to see the signs) and directed them to indicate whether or not the actual sign was the same or different. Participants were then allowed to activate their wipers and adjust the settings at their discretion. The experimenter noted what setting was used on every trial. The cart was advanced toward the participant's vehicle at a comfortable walking pace. When the participant made a choice, ground measurements were used to determine the distance between the sign and the car.

Three of the six trials were match trials and three were mismatch trials. The six trials occurred in one of two alternating orders.

Results.

We have complete data from 22 younger and 20 older adults.

Accuracy

ANOVA showed a significant effect of sign type only, $F(1, 40) = 14.7$, $p < .01$. Accuracy (proportion correct) was higher on symbol signs ($M = .86$, $SE = .031$) than on word signs ($M = .71$, $SE = .032$). This result replicates the finding in the clear weather day conditions for the field component in study 3.

Distance

Given the low number of trials presented to each driver, in order to preserve power we analyzed distance data for both accurate and inaccurate trials. Again, the only significant effect was type of sign, $F(1, 40) = 16.3$, $p < .01$. Symbol signs were identified at a distance of 458 feet ($SE = 34.2$) whereas word signs were perceived from 389.4 feet ($SE = 32.5$). This replicates the advantage for symbol warning signs shown in Study 3 under clear weather conditions. However, the distances were reduced considerably compared to that study where signs were comprehended at 831 feet for symbol, and 505 feet for word signs.

Conclusions.

Surprisingly, age was not a significant factor in the comprehension of signs under adverse weather conditions (rain simulation). Young and older drivers achieved roughly equivalent performance (accuracy, distance). However, adverse weather conditions (intense rain) seriously impeded performance compared to the dry weather conditions in study 3. Symbol signs were more impaired than word signs in terms of distance reduction, but still had a large advantage for comprehension. Thus, we continue to recommend that warning signs be displayed as symbol rather than word message signs. They show a clear advantage in both good and poor weather conditions.

Caveats

We do not know whether rain combined with darkness might not selectively disadvantage older drivers. Also, legibility assessed with a comprehension task may not predict driver behavior though we would expect that the advantage shown for comprehending a symbol sign will translate into more effective action on the road.

Benefit of the Project

This project has provided relevant data to aid the formulation of policy and recommendations for the Safe Mobility for Life Program. Some of the findings with relevant policy implications are:

1) The finding of a minor or negligible advantage of more expensive microprismatic fluorescent sheeting for warning signs compared to diamond grade sheeting. We were unable to demonstrate a robust advantage for fluorescent sheeting for sign legibility in Study 1.
2) All of our studies have pointed to the importance of enrolling older drivers in research about traffic safety because they are the most sensitive to factors influencing the legibility of traffic signs and signals. They are also the most sensitive to factors influencing intersection traversal both as motorists and pedestrians.
3) We showed a reasonably consistent legibility advantage for symbol warning signs over word warning signs (day, night, and heavy rain conditions).
4) Uppercase lettering for street signs improved legibility over mixed case lettering, particularly for older drivers.
5) Given that we had a mixed pattern of prediction from lab to field performance, we have demonstrated that laboratory studies should be supplemented by field studies. Field studies are necessarily more expensive to run than lab studies but it is evident that they are a necessary component in obtaining valid data for policy formulation and implementation.
6) We have also provided relevant data to help guide future research on driver and pedestrian safety. Findings that warrant follow up research include:
    a. Assessing high versus low beam headlamp settings for their effect on legibility of fluorescent versus standard sheeting for warning signs.
    b. Assessing the utility of contractions for long street names to improve street name legibility.
    c. Assessing the structuring of messages for portable changeable message signs.
    d. Assessing different pedestrian signals and their timing to ensure that older pedestrians cross intersections safely.

Specific Recommendations

Based on the findings in the project, we can provide the following recommendations.

Study 1.

We do not recommend adopting fluorescent sheeted warning signs if their expense is significantly greater than that for standard sheeting. More research is needed to establish the relative efficacy of that sheeting for different types of signs under both low and high beam settings.

Given that increased driver age magnified the effects of other variables (such as distance in the lab task), we recommend that older drivers be consistently incorporated into studies of traffic sign perception. They are the most sensitive to varying sign attributes.

Study 2.

We recommend that number of letters in street signs be minimized where possible, perhaps through the use of contractions, though perception of contractions needs further empirical investigation. We also recommend that uppercase lettering be used rather than mixed case for street names to improve decision making time, particularly for older drivers.

Study 3.

We strongly recommend the use of symbol warning signs in place of word warning signs based on their legibility advantage.

Study 4.

We recommend that research be conducted into ways to enhance the legibility of PCMS messages to reduce the comprehension gap between older and younger drivers.

Study 5.

We recommend that further research, particularly field studies, be conducted into the decision making process of pedestrians, particularly older ones, to promote safer crossing behavior.

Study 6.

We recommend that those setting timing intervals for pedestrian crossings should assume slower walking speeds than 4 ft/s particularly for intersections that receive above average visits from older pedestrians to be sure that they have adequate time to cross safely.

Study 7.

Given that the literature indicates that many pedestrian crashes occur at night, we recommend that further research be carried out at night with field studies examining how driver velocity profiles change in the presence of pedestrians varying in conspicuity.

Study 8.

We strongly recommend the adoption of symbol warning signs compared to word warning signs for providing the most advanced warning to drivers under adverse weather (rain) conditions.

References

MUTCD (2003, with revisions 1 and 2). Accessed from http://mutcd.fhwa.dot.gov/ 7/3/09.

Anders, R. L. (2000). On-road investigation of fluorescent sign colors to improve conspicuity. Unpublished Masters thesis, Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

Bédard, M., Weaver, B., Darzins, P., & Porter, M. M. (2008). Predicting driving performance in older adults: We are not there yet! Traffic Injury Prevention, 9, 336-341.

Botha, J. L., Zabyshny, A. A., Day, J. E., Northouse, R. L., Rodriguez, J. O., & Nix, T. L. (2002). Pedestrian countdown signals: An experimental evaluation. Vol. 1. Final Report to the California Traffic Control Devices Committee.

Charlton, S. G. (2006). Conspicuity, memorability, comprehension, and priming in road hazard warning signs. *Accident Analysis & Prevention, 38*, 496-506.

Charness, N., & Bosman, E. A. (1990). Human factors and design for older adults. In J. E. Birren & K. W. Schaie (Eds.), Handbook of the psychology of aging (3rd ed.). (pp. 446-463). San Diego: Academic Press.

Chrysler, S. T., Carlson, P. J., & Hawkins, H. G. (2002). Nightime legibility of traffic signs as a function of font, color, and retroreflective sheeting. Paper presented at the 82nd Annual Meeting of the Transportation Research Board. Paper Number 03-4113.

Chrysler, S., Stackhouse, S., Tranchida, D., & Arthur, E. (2001). Improving street name sign legibility for older drivers. Proceedings of the Human Factors and Ergonomics Society, 45th Annual Meeting.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd edition). Hillsdale, NJ: Erlbaum.

Dewar, R., Kline, D., Scheiber, F., & Swanson, A. (1997). Symbol signing design for older drivers. FHWA-RD-94-069.

Dunbar, G., Holland, C. A., & Maylor, E. A. (2004). Road Safety Research Report No. 37. Older pedestrians: A critical review of the literature. Department for Transport, Great Minster House, 76 Marsham Street, London SW1P 4DR, Tel: 020 7944 8300, Internet service: www.dft.gov.uk.

Eccles, K.A., & Hummer, J.E. (2001). Safety effects of fluorescent yellow warning signs at hazardous sites in daylight. *Proceedings from the TRB 80th Annual Meeting*, January 7-11, 2001, Washington D.C.

Edwards, C. J., Creaser, J. I., Caird, J. K., Lamsdale, A. M., & Chisholm, S. L. (2002). Older and younger driver performance at complex intersections: Implications for using perception-response time and driving simulation. *Proceedings of the Second International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*. Park City, Utah, USA.

Emo, A. K., & Do, A. H. (2008). Evaluating and validating a new pedestrian exposure to risk metric. *Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting, 52,* 1845-1849.

Evans, L. (2004). Traffic safety. Bloomfield, Michigan: Science Serving Society.

Finnis, K. K., & Walton, D. (2008). Field observations to determine the influence of population size, location and individual factors on pedestrian walking speeds. *Ergonomics, 51*, 827-842.

Fisher, J., (1992). Testing the effect of road traffic signs' informational value on driver behavior. *Human Factors ,34*, 231–237.

Fisk, A. D., Rogers, W. A., Charness, N., Czaja, S. J., & Sharit, J. (2009). Designing for older adults: Principles and creative human factors approaches (2nd Ed.). Boca Raton: CRC Press.

Forbes, T. W., Moscowitz, K., & Morgan, G. (1950). A comparison of lower case and capital letters for highway signs. *Proceedings of the 30th Annual Meeting of the Highway Research Board*, Washington, D.C.

Garvey, P. M., Pietrucha, M. T., & Meeker, D. T. (1997). Effects of font and capitalization on legibility of guide signs. Transportation Research Record, 1605, 73-79.

Geruschat, D. R., & Hassan, S. E. (2005). Driver behavior in yielding to sighted and blind pedestrians at roundabouts. J*ournal of Visual Impairment and Blindness, 99*, 286-302.

Guerrier, J. H. and S.-H. Fu. Elder Roadway User Program Test Sections and Effectiveness Study. Final Report Task 4, University of Miami Project No. 669535, Florida Department of Transportation Research Center, Tallahassee, FL, October 2002.

Hakamies-Blomqvist, L., & Wahlstrom, B. (1998). Why do older drivers give up driving? *Accident Analysis and Prevention, 30*, 305-312.

Holick, A. J., & Carlson, P. J. (2003). Nighttime Guide Sign Legibility for Microprismatic Clearview Legend on High Intensity Background. FHWA/TX-04/0-1796-4. Texas Department of Transportation Research and Technology Implementation Office, P. O. Box 5080, Austin, Texas 78763-5080.

Huang, H., & Zegeer, C. (2000). The effects of pedestrian countdown signals in Lake Buena Vista. Report for Florida Department of Transportation. Accessed 2/13/08 from http://www.dot.state.fl.us/Safety/ped_bike/handbooks_and_research/research/CNT-REPT.pdf

Jenssen, G.D., Moen, T., Brekke, B., Augdal, A. & Sjohaug, K. (1996). Visual performance of fluorescent retroreflective traffic control devices. Technical Report STF22-A96606. Trondheim, Norway: SINTEF Transport Engineering.

Kim, K., Brunner, I. M., & Yamashita, E. (2008). Modeling violation of Hawaii's crosswalk law. *Accident Analysis and Prevention, 40,* 894–904.

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

Knoblauch, R. L., Pietrucha, M. T., & Nitzburg, M. (1996). Field studies of pedestrian walking speed and start-up time. *Transportation Research Record 1538*.

Kohn, L. T., Corrigan, J. M., & Donaldson, M. S. (Eds). (2000). To err is human. Building a safer health system. Washington, DC: National Academy Press. (http://www.nap.edu/catalog/9728.html)

Langham, M., & Moberly, N. (2003). Pedestrian conspicuity research: A review. *Ergonomics, 46*, 345-363.

Manual on Uniform Traffic Control Devices, MUTCD (2003, with revisions 1 and 2). Accessed from http://mutcd.fhwa.dot.gov/ 7/3/09.

Morrow, D. G. & Rogers, W. A. (2008). Environmental support: An integrative framework. *Human Factors, 50*, 589-613.

Nasvadi, G. E., & Vavrik, J. (2007). Crash risk of older drivers after attending a mature driver education program. *Accident Analysis and Prevention, 39*, 1073-1079.

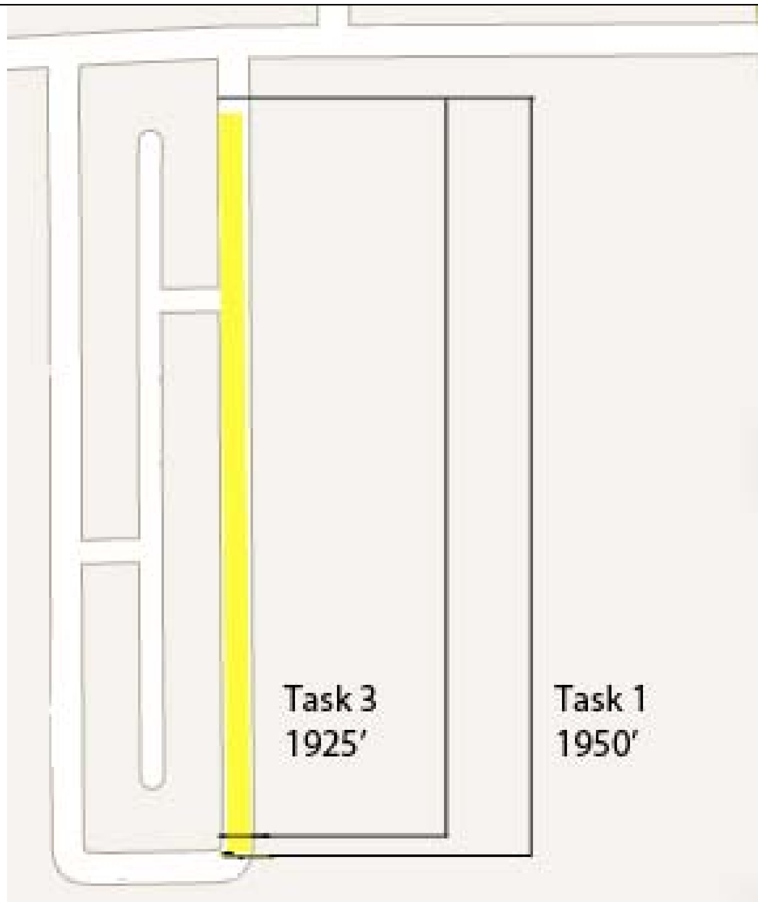
Oxley, J., Fildes, B. N., & Dewar, R. E. (2004). Safety of older pedestrians. In

Transportation in an aging society: A decade of experience (pp. 167-191). Transportation Research Board Conference proceedings 27, Washington, DC.
Rothe, J. P. (1990). *The safety of elderly drivers: Yesterday's young in today's traffic*. New Brunswick, NJ: Transaction Publishers.
Schattler, K. L., Wakim, J. G., Datta, T. K., & McAvoy, D. (2007). Evaluation of Pedestrian and Driver Behaviors at Countdown Pedestrian Signals in Peoria, Illinois. *Transportation Research Record: Journal of the Transportation Research Board, 2002*, 98-106.
Schnell, T., Bentley, K. Hayes, E., & Rick, M. (2001). Legibility distances of fluorescent traffic signs and their normal color counterparts. In Transportation Research Record 1754, National Research Council, Washington, D.C.
Sharit, J. (2006). Human error. In G. Salvendy (Ed.). Handbook of Human Factors and Ergonomics (3rd Ed.) (pp. 708-760). Hoboken, NJ: Wiley.
Singer, J. P., & Lerner, N. D. (2005). Countdown pedestrian signals: A comparison of alternative pedestrian change interval displays. Final Report to Federal Highway Administration. Accessed 6/24/09 from http://www.atssa.com/galleries/default-file/Ped_Countdown_Report.pdf
Smiley, A., Courage, C., Smahel, T., Fitch, G., & Currie, M. (2001). Letter height for street names: An on road study. *Proceedings of the SELF-ACE 2001 Conference – Ergonomics for changing work, 4*, 197-202.
Steenbekkers, L. P. A., & van Beijsterveldt, C. E. M. (Eds.) (1998). Design-relevant characteristics of ageing users. Delft, The Netherlands: Delft University Press
Stollof, E. R., McGee, H., & Eccles, K. A. (2007). Pedestrian signal safety for older persons. AAA Foundation for traffic safety, 607 14th Street, NW, Suite 201, Washington, DC 20005. *www.aaafoundation.org/pdf/PEDsigtiming.pdf, downloaded 6/7/09.*
Transportation in an aging society: A decade of experience. (2004). Transportation Research Board Conference proceedings 27, Washington, DC.
Wood, J. M., Tyrrell, R. A., & Carberry, T. P. (2005). Limitation in drivers' ability to recognize pedestrians at night. *Human Factors, 47*, 644-653.
Zwahlen, H. T., & Schnell, T. (1999). Legibility of traffic sign text and symbols. Paper No. 980284, presented at the 77th Annual Meeting of the Transportation Research Board, Washington, DC, January 11-15, 1998. Published in *Transportation Research Record, 1692*, 142-151.

Appendix A. Field Testing Locations

Location 1 of field testing, “Broadmoor Estates”.

Track:



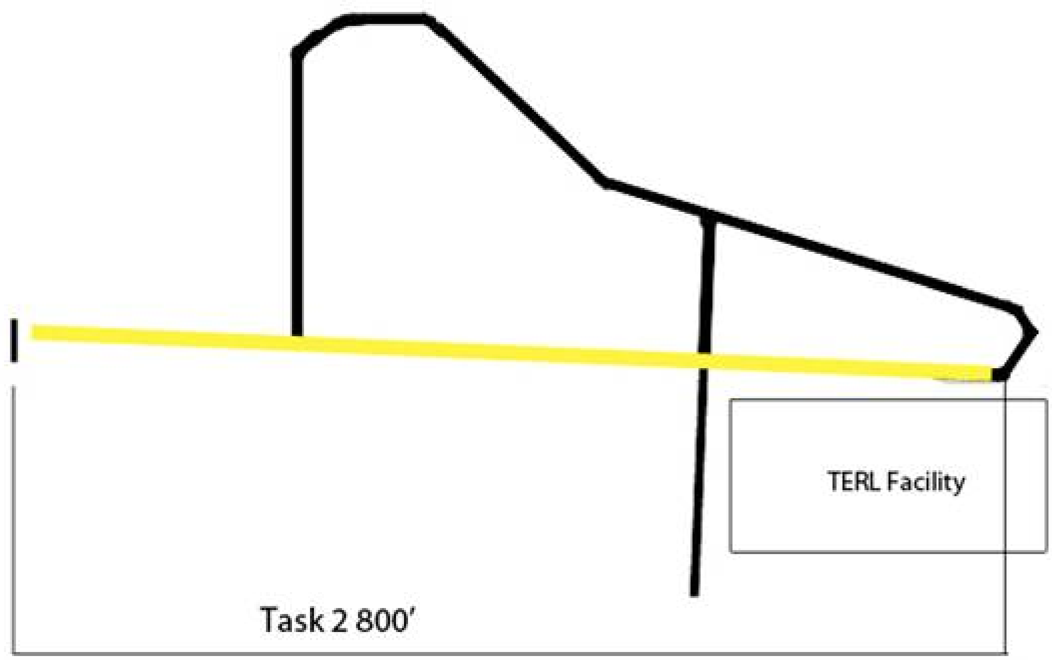
Description. The facility was located on city property behind fences. The track was a straight road spanning approximately 2000’. For our purposes we used 1,950 feet for task 1, and 1,925 feet for task 3. The remaining footage was left for experimental equipment and an area for participants to turn around in.

Photo of “Broadmoor Estates” track:

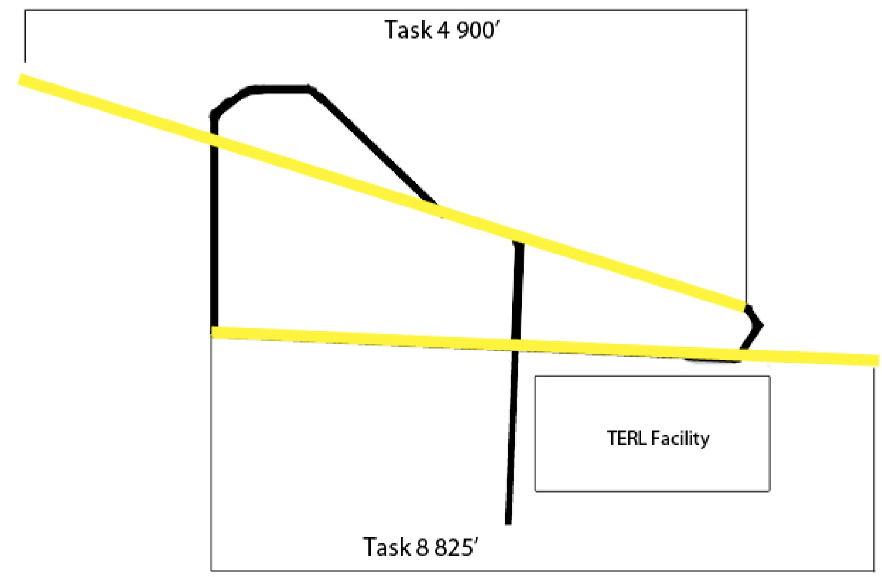
Description. This photo shows from near starting point to the end of the track during the experiment with a participant. The orange ‘dot’ at the end is actually a warning sign used for the testing. This was the same set up for both task 1 and task 3.

Location 2 of field testing, “TERL” (Traffic Engineering and Research Lab).

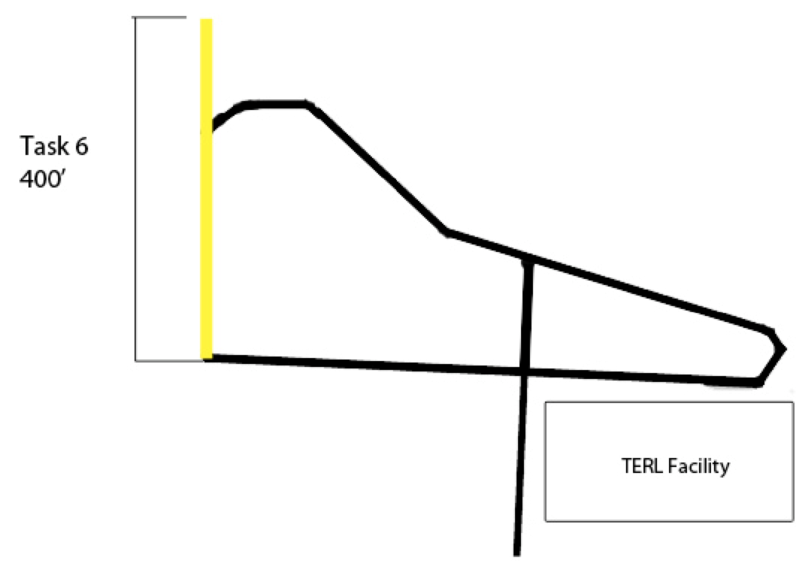
The following maps show the layouts used for each task within the project at the TERL Facility. The maps show the ‘backyard’ of the facility and the roads in existence. Lines outside of these roads lead into the yard areas and were included in the tasks when notated.



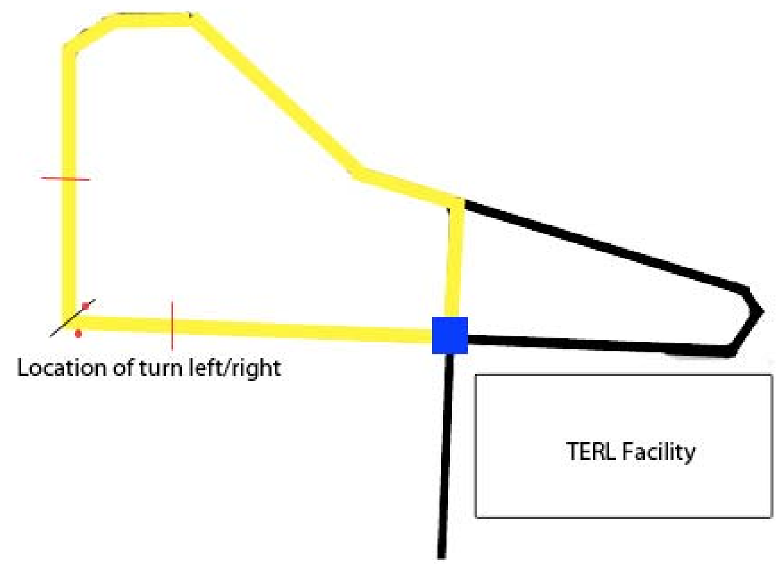
Description. The yellow indicates the area used for Task 2. A total of 800 feet was used for this task.



Description. These yellow lines denote the areas used for tasks 4 and 8. Task 4 used 900 feet while task 8 used only 825 feet.



Description. The yellow line denotes the area used on task 6. Task 6 used 400 feet.



Description. The yellow lines in this map denote the short track used for the route taken by the participant. Participants would travel either clockwise or counter-clockwise to achieve either a right-hand turn or a left-hand turn. The red lines indicate the area of radar data collection. The blue square indicates the area in which participants waited in between trials. The two red dots indicate the areas in which the pedestrians started from. The top dot was used during right turn trials, and the bottom dot indicates the area used during left hand turns.

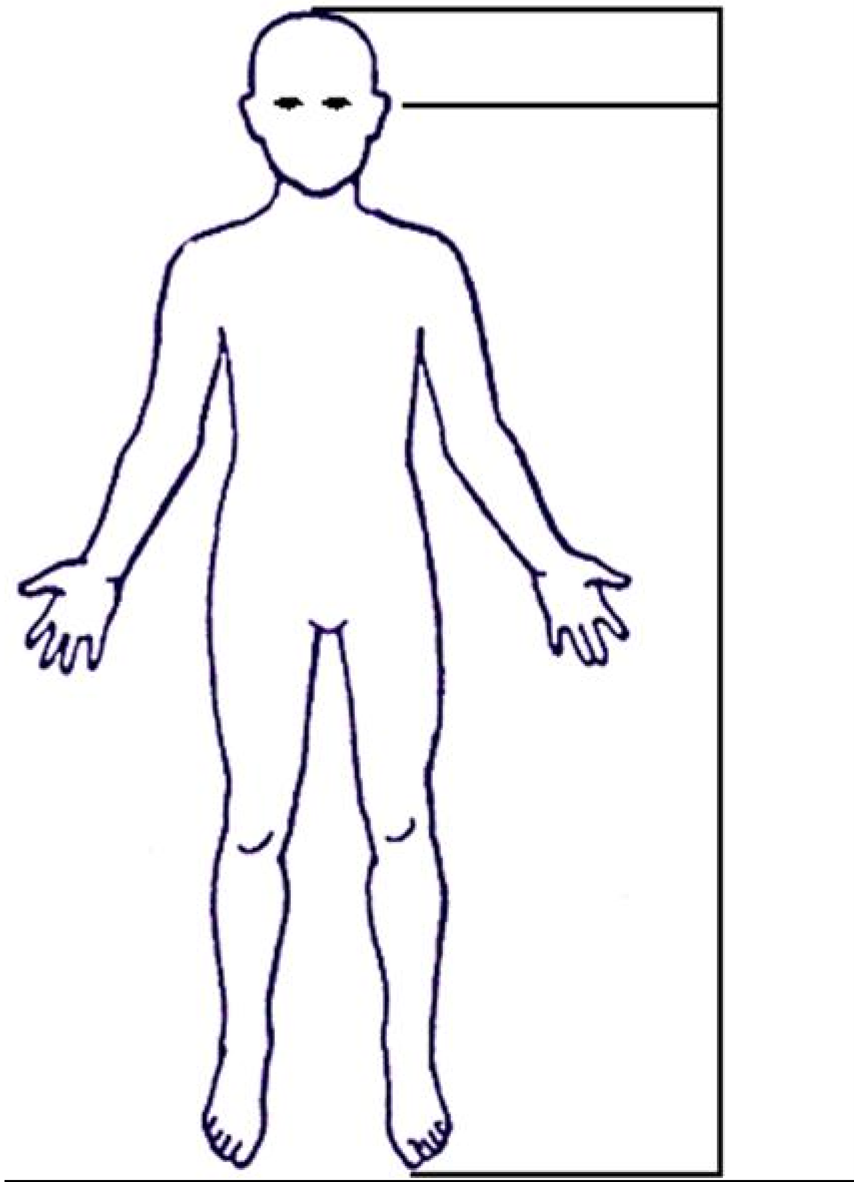
Photos of the TERL 'backyard':

Description. The picture above depicts a mid-track view of the set up for Task 2. The Sign apparatus can be seen underneath the three red lights. The distance to the sign in this picture is approximately 600 feet.

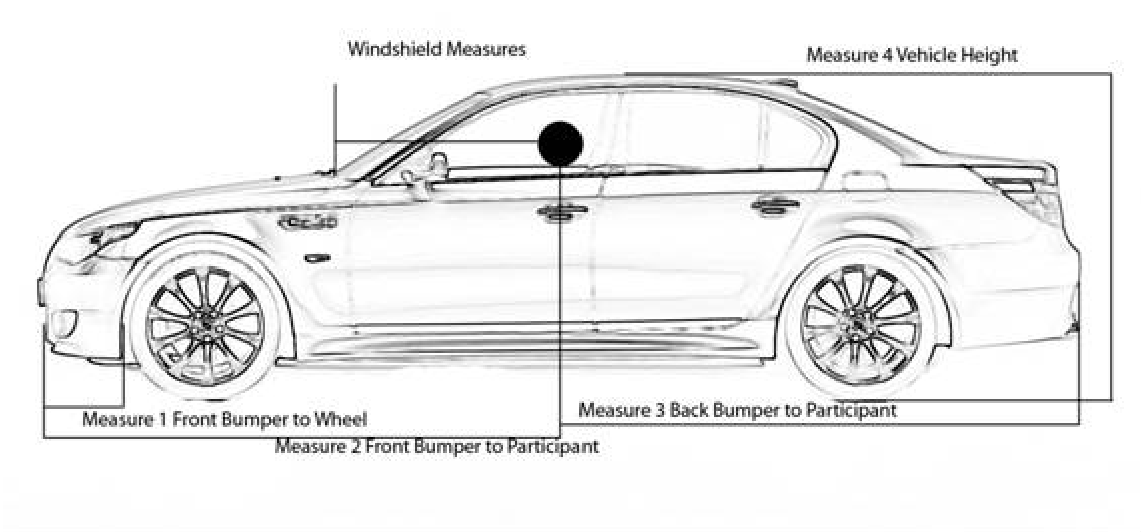
Description. Photo above depicts the view participants saw for task 4. The viewing distance to the sign is 900 feet. Please note, the PCMS sign pictured does not appear to have all letters present. This is a photographic anomaly; the sign is displaying all letters.

Description. This picture is a view of the task 8 track that signs took. This is looking through the weather-mimicking device. This view shows all 825 feet.

Participant and car measurements:



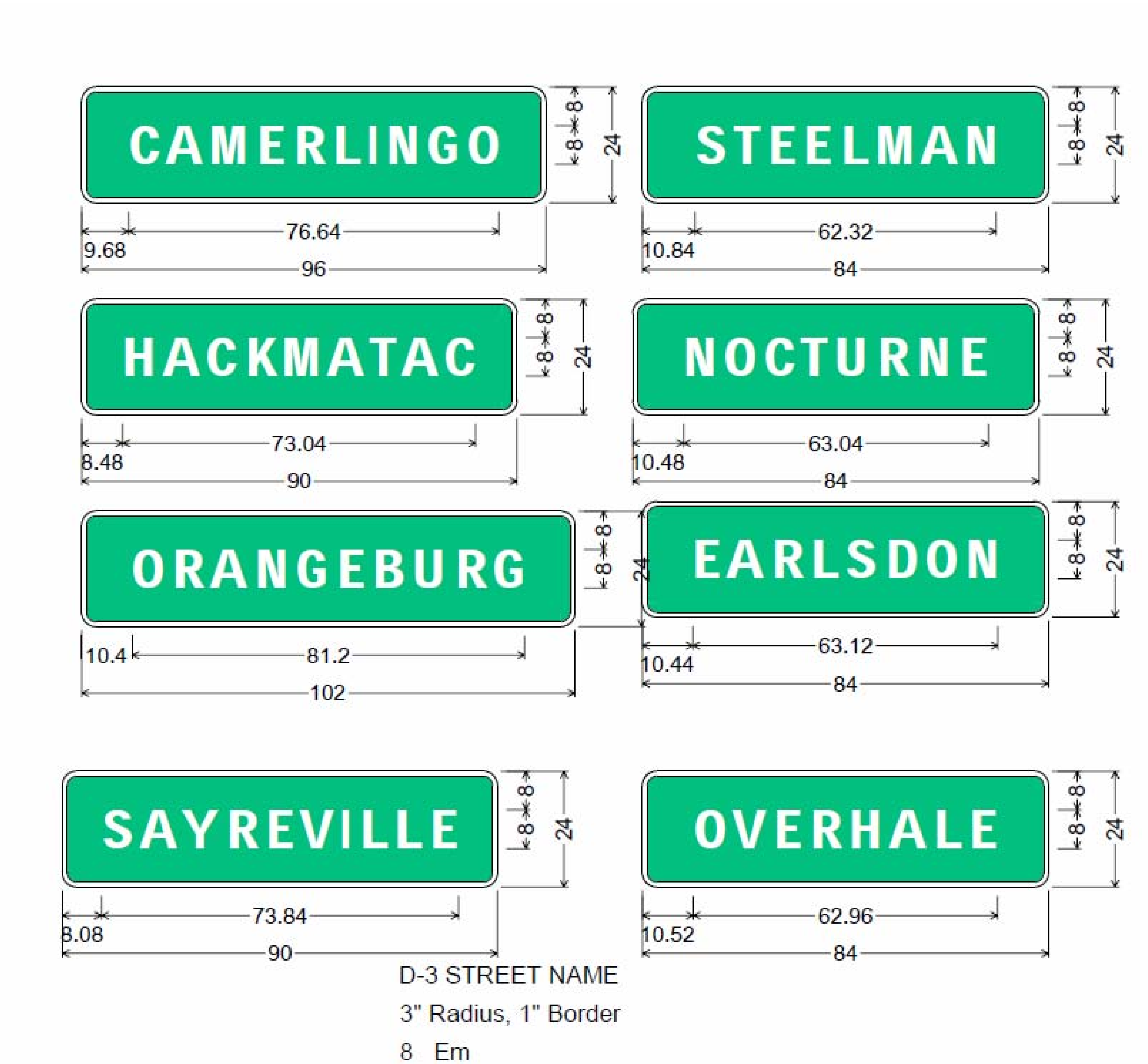
Description. This diagram depicts the measurements gathered on participants outside of their vehicles. Total height and height to eyes was attained.



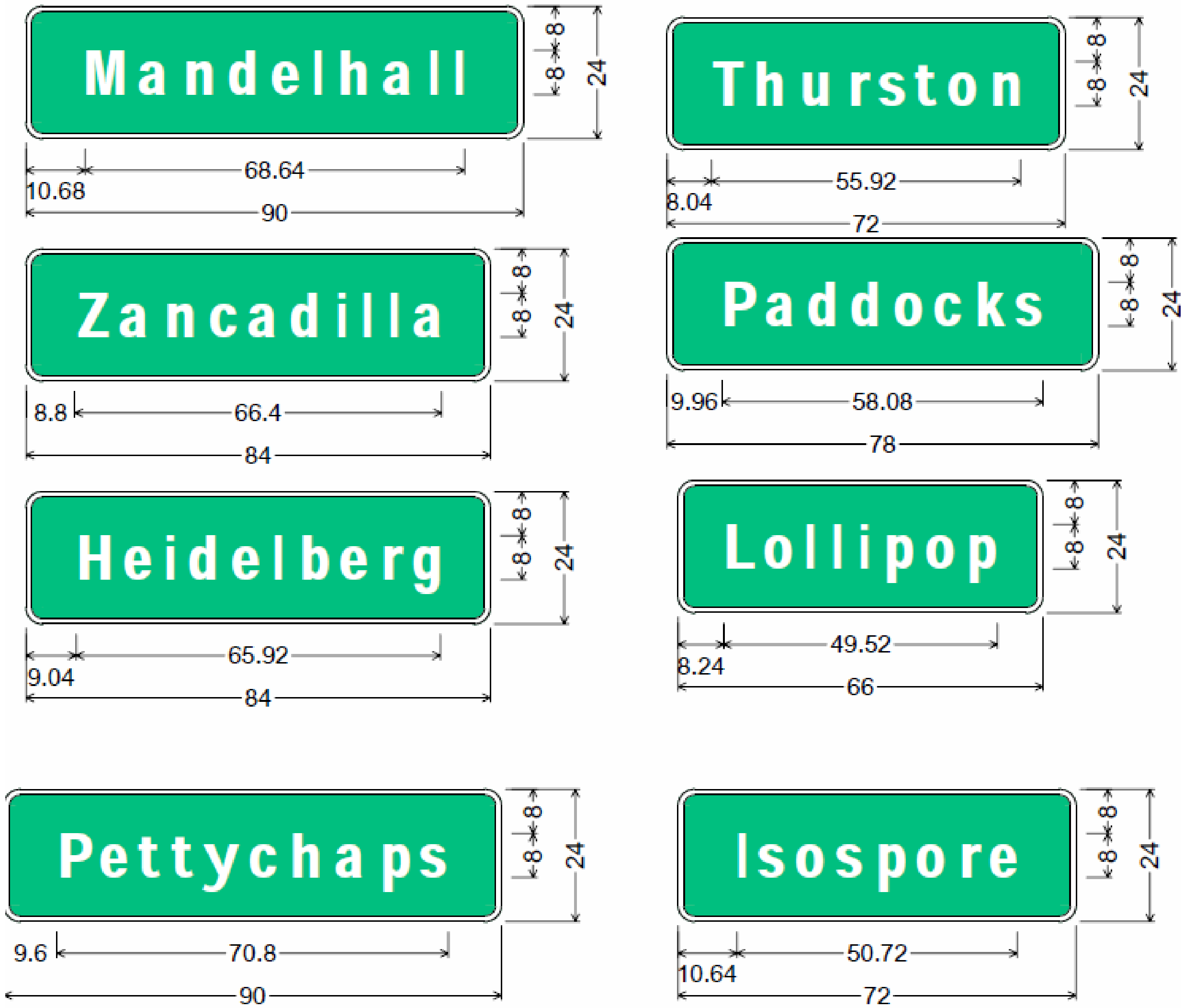
Description. This diagram shows the measurements attained on the cars and the participants in the cars for all tasks except for tasks 5 and 6.

Appendix B. Testing Equipment

Task 2 field street signs: Uppercase, then mixed case.



Note. Hackmatac was corrected to Hackmatack on the actual sign.



D-3 STREET NAME

3" Radius, .1" Border

8/6 Em